AD_____

Award Number: DAMD17-02-1-0082

TITLE: The Hinge Region as a Key Regulatory Element of Androgen Receptor Dimerization, DNA Binding, and Transactivation

PRINCIPAL INVESTIGATOR: Frank A. Claessens, Ph.D.

CONTRACTING ORGANIZATION: K.U. Leuven Research and Development
B-3000 Leuven
Belbium

REPORT DATE: May 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2005	Annual (1 Apr 2004 - 31 Mar 2005)

4. TITLE AND SUBTITLE
The Hinge Region as a Key Regulatory Element of Androgen Receptor Dimerization, DNA Binding, and Transactivation

5. FUNDING NUMBERS
DAMD17-02-1-0082

6. AUTHOR(S)
Frank A. Claessens, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
K.U. Leuven Research and Development
B-3000 Leuven
Belbium

E-Mail: Frank.claessens@med.kuleuven.ac.be

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The androgen receptor binds inverted repeats as well as direct repeats of the 5'-AGAACA-3' core element. The inverted repeats are called classic AREs, the direct repeats selective AREs. We found recently that the progesterone receptor is also able to transactivate through direct repeat elements, but since progesterone concentration in male serum is low, and since the ARE are found in androgen-responsive genes, the elements can still be called AR-selective.

In collaboration with the group of Daniel Gewirth (DAMD17-01-1-0050) we were able to solve the crystal structure of the AR DNA-binding domain bound to a direct repeat. The data indicate a stronger dimerization interface between the two AR protomers bound to the direct repeat in a unexpected head-to-head conformation. Our functional analyses after introduction of mutations in AR as well as GR do not corroborate the structural data. Further mutation analysis of the CTE is being done.

A carboxyterminal extension of the DNA binding domain of the AR is known to be involved in DNA binding (unfortunately its structure remains unsolved), in nuclear localization and in transactivation control. Surprisingly, deletions which affect DNA binding negatively have a positive effect on transactivation when tested in full size receptors. We have analyzed the effect of the deletion on the two activation functions of the AR. In conclusion, we observed an increased stability of the AR upon deletion of the CTE, but this is not perfectly correlated with an increased androgen response. Point mutations in the hinge can partially mimic the effect of e deletion. The lysines in the hinge seem to be major signal input residues, but their exact roles are specific for each lysine (degradation control versus transactivation control and/or nuclear localization). Further research is aiming to discriminate these.

14. SUBJECT TERMS
Androgen receptor, selectivity, DNA binding, transactivation

15. NUMBER OF PAGES
43

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

THE HINGE REGION AS A KEY REGULATORY ELEMENT OF ANDROGEN RECEPTOR DIMERIZATION, DNA BINDING AND TRANSACTIVATION

INTRODUCTION

Androgens play specific key roles in normal male development as well as in many physiological processes in both sexes. In addition, they are involved in the development and progression of prostate cancer, a pathology that affects a high proportion of the aging men. The androgen receptor is a ligand-dependent transcription factor. Together with the receptors for glucocorticoids, progestagens and mineralocorticoids it binds bipartite hormone response elements organized as inverted repeats of 5'-AGAACA-3'-like core elements separated by three nucleotides. We discovered that the androgen receptor can bind direct repeats of the same core elements as well (partly reviewed in Claessens *et al.* 2001). This feature distinguishes this receptor from the glucocorticoid receptor. The inverted repeats are called classic AREs, while the direct repeats will be called selective AREs.

Structural determination based on X-ray data on co-crystals of the DNA-binding domain of the androgen receptor (AR-DBD) with a selective ARE led to a structure depicted in figure 1A. Two zinc coordinating modules that constitute the receptors DNA-binding domain, are involved in the recognition of classical AREs, but for the high affinity binding to selective AREs, a short amino acid sequence at the carboxyterminus of the DNA binding domain is needed (figure 1B). This extension is not only involved in DNA binding, but is also involved in nuclear translocation. In addition, the hinge region is reported to control the transactivating properties of the androgen receptor (Moilanen *et al.* 1997, Wang *et al.* 2001). In this study, we are unraveling the structure-function relationships in the two zinc coordinating modules and part of the hinge region (reviewed in Claessens and Gewirth, 2004).

BODY

We have translated the structural data derived from the X-ray data (Shaffer *et al.* 2004) in part I. In earlier work (see 2004 report), we observed that a deletion of the hinge region results in an androgen receptor which is more active compared to the wild type AR (Figure 1C). This observation is studied further in part II.

Part I. Co-Crystals of an AR-DBD fragment with a direct repeat ARE and translation of the structural data

Co-crystals have been obtained, and the resulting structural data have been reported by Shaffer *et al.* 2004 (see also Figure 1A, appendix 1 and the 2004 report). The co-crystals contained AR-DBD dimers bound to a direct repeat of the 5'-TGTTCT-3' hexamer separated by a three nucleotide spacer. Although earlier data indicated that the AR would dimerise on such sequences in a head-to-tail conformation, the crystal data showed a head-to-head AR-DBD dimer. The AR dimerization interface involved is very similar to that described for the GR-DBD. In the case of the AR, additional hydrogen bonds are formed between a Threonine (an Isoleucine in the GR) of one monomer and a keto-function in the peptide backbone of the other DBD (and vice versa). In addition, at the position of a Glycine in the GR-DBD, the AR-DBD has a Serine, which increases the contact surface and which forms an additional hydrogen bond with its counterpart in the other monomer.
Based on these observations, we concluded that the AR can recognise direct repeat elements because of its stronger dimerisation. The GR, on the other hand, can not dimerise on such sequences. In this way, direct 5'-TGTTCT-3' repeats are androgen selective.

I.A. Translation of crystal data

We have tested this hypothesis by exchanging the Threonine and Serine in the AR by Isoleucine and Glycine respectively. Much to our surprise, in transient transfection experiments, none of the mutations lead to a change in specificity of the receptors (Figure 2 A). Hence, the reduction of the number of hydrogen bonds did not affect the transactivation by the AR via selective AREs. The establishment of these bonds in the GR, did not make it activate through the selective AREs, either.

These same mutations were introduced in isolated DNA-binding domains of AR and GR. The DBDs were cut from the GST-fusion used to purify them from *E.coli* extracts. In band shift assays, the selectivity was not changed, even when the double mutants were tested: selective AREs are only bound by AR-derived DBDs, while classical AREs are bound by both AR and GR-derived DBDs with or without the indicated mutations (Figure 2B). Clearly, **the stronger dimerization surface mediated by the two AR-specific residues in the so-called D-box is not sufficient to explain the changes in specificity between AR and GR**.

From very early swapping experiments we learned that, besides the second zinc coordinating module of the DBD, a short 12 amino acid long C-terminal extension of this fragment, called CTE, is involved in the recognition of the selective AREs (Schoenmakers et al. 1999). In this CTE, two residues were relevant to the binding of selective AREs, but not to classical AREs (Schoenmakers et al. 2000). We are now exchanging these residues between AR and GR, as well as between AR- and GR-DBDs. The effect of these mutations will be studied in transient transfections and band shift assays, respectively. Much to our surprise, the development of mutations in these residues (this fragment of the AR cDNA) is proving very difficult. Indeed, where the introduction of mutations very near to these residues was succesful earlier (see 2003 and 2003 reports), several attempts with divergent protocols have only slowly been developing. However, an alternative PCR-based mutagenesis is now being used to introduce mutations in this part of the hinge region.

I.B. Selectivity of the androgen responses

Testing the AR for selectivity towards AREs was done in transient transfection experiments which enabled the use of cell lines that express no endogenous steroid receptors. A selective ARE was defined as such when the hormone responses obtained by co-transfection with an AR expression plasmid was much higher than that obtained by co-transfection with a GR expression plasmid (reviewed in Claessens et al. 2001). This led first to the description of the PB-ARE-2 (Claessens et al. 1996), later of scARE and slpHRE2 as selective AREs (Verrijdt et al. 2000). The groups of Trapman and Haendler have added SARG-AR (Steketee et al. 2004) and the pem ARE (Barbelescu et al. 2001) to this list.

More recently, we have tested the responsiveness of these elements to progestagens. Based on the higher homology with between PR and GR versus PR and AR, we initially thought this would be negative. However, since the earlier experiments showed that the CTE might be

more important in selectivity, and since the first mutation analyses of the D-box did not corroborate with the Crystal data, we turned our attention to the CTE. We remarked that the AR-specific Glycine (position 627) is conserved in the PR, while for the Leucine at position 634 in AR, a Phenylalanine is present in the PR (Figure 3A).
In transient co-transfection experiments, we indeed observed that the PRB, like AR, is able to transactivate through the C3(1) ARE as well as through the PB-ARE-2 (Figure 3B). The hormone concentration needed to observe a response is higher than the reported male serum concentration (Figure 3C). Since progesterone is low in male serum, and since the PR will normally not be expressed and activated under conditions where the androgen-responsive genes are active, the AREs can still be called selective. A clear function of progesterone in male mice is in the control of behaviour towards offspring (Schneider et al. 2003). Alternatively, direct repeat-like elements could result in progesterone-specific gene regulation in female mice, in tissues where androgens are low. The selectivity of the androgen respons via selective AREs will be tested in ARE-based reporter mice (other project).
Mutation analysis will now have to indicate whether the Glycine at position 627 and the Leucine at 634 are indeed responsible for the AR-selectivity.

I.C. Fluorescence Resonance Energy Transfer (FRET)

The aim of this experiment is to analyse conformational differences of the AR dimer binding to an AR-selective versus a classic ARE in solution. DNA oligonucleotidess that contain the respective AR binding sites have been internally labeled with fluoresceine and can serve as FRET donors. The FRET-acceptor is the AR-DBD labeled with QSY 35 (Molecular Probes) at the sulfhydryl group of a specific cysteine (not involved in the coordination of zinc). Depending on the conformation of the proteins bound to the oligonucleotides, we expect differences in fluorescence-quenching. These differences can be used to calculate the distances between the donor-acceptor pair giving us important information about the orientation of the AR-dimer on the different types of binding sites. These experiments are being performed in collaboration with the group of Dr. Engelborghs (Laboratory for Biomolecular Dynamics, KULeuven).

II. Structure-function relationships within the hinge region of the AR

Part of this work has been described in the earlier reports. In short, we delineated the minimal region of the hinge region necessary for the 'superactivation' of the AR (Figure 1B and C).

II.A. Studies in yeast

As reported earlier, the deletion of the hinge region has no implications on the activity of a Gal4 DBD-fused AR-LBD when tested in yeast. This prevented us from setting up a yeast screen for activating mutations in the hinge region. In collaboration with Dr. Ceralin (Laboratoire de Cancerologie Experimentale et de Radiobiologie-EA/ULP 3430, IRCAD, Strasbourg, France), we will analyse the effect of some of the mutations in yeast strains which contain ARE-based reporters (Ceraline et al. 2003). In these yeast, we will analyse the transactivating properties of the AR or AR-fragments which are recruited to DNA via their own DNA-binding domain. This is important in view of the role of the different functions of the CTE in DNA-binding, nuclear localization and transcativation control.

II.B. Role of the hinge region on activity of AF2

The deletion of the hinge region was shown to have an effect on the activity of the AF1 (NTD-DBD, will be discussed in part II.G), as well as on the isolated AF2. The latter effect is more pronounced when the transactivation is tested on a GAL4-responsive reporter (Figure 4A). It is unclear at the moment what explains the dependence of this difference on the presence of the heterologous DBD.

Because the AF2 of nuclear receptors is know to recruit p160s, we have analyzed the effect of the deletion of the hinge (Δ1) on the recruitment of the SRC1A fragment with highest affinity for the AR-LBD. Indeed, the central nuclear receptor interacting region has only low affinity for the AR-LBD, while the LxxLL-motif in the carboxyterminal end of the splice variant SRC1A has a high affinity (Bevan *et al.* 1999). In double hybrid experiments, the deletion of the hinge results in a decreased binding of the SRC1a fragment (Figure 4B). However, it should be noted that in the case of the AR, a strong interaction between the aminoterminal domain and the ligand binding domain (the N/C intercation) excludes a recruitment of p160 coactivators to the LBD via LxxLL motifs (Bevan *et al.* 1999, Alen *et al.* 1999). Indeed, it is the AR-NTD that recruits the

p160s through a glutamine-rich region (Christiaens *et al.* 2003). We will now analyse the possible implication of the hinge in the interaction between this region and the amino-terminal domain of the AR.

We conclude that the deletion of the hinge affects the activity of AF2. This contrasts with the observation that the SRC1a coactivator fragment is less efficiently recruited.

II.C. Role of the hinge region in the N/C interactions

The N/C interaction in the AR depends on a 'FQNLF' motif in the aminoterminal domain and the integrity of the LBD (He and Wilson 2002; Dubbink et al. 2004). When this motif is deleted, the function of the AR is largely impaired (Callewaert *et al.* 2004). This is also true in the absence of the hinge region, since the transactivation of AR with deletion of both the hinge and the FQNLF is reduced approximately 5-fold (figure 5A). In the AR in which the FQNLF is deleted, the removal of the hinge still has a strong potentiating effect, indicating that the two mechanisms are independent.

Not surprisingly, in a double hybrid assay, the deletion of the FQNLF resulted in a dramatic drop of the N/C interactions whether the hinge region was present or not (figure 5B).

We conclude that the higher potency of the AR without hinge region is not due to an enhanced N/C interaction.

II.D. Role of a putative PEST sequence and a phosphorylation site

The presence of a putative PEST sequence within the hinge region, next to the Δ11 motif indicates a possible communication between these different functions (Figure 6A).

We mutated the PEST sequence, as well as two flanking Lysine residues (638 and 658) which would be the presumed targets for poly-ubiquitination. The deletion of the PEST sequence resulted in a less active AR when tested on the MMTV- or the C3(1)-based reporter, in COS 7 cells and in HeLa cells (Figure 6B). This is in contrast with the potentiating effect of the Δ11 deletion (Figure 1C). From the two mutations of Lysines 638 and 658 into Arginines, only the latter had an effect similar to that of the PEST sequence deletion (Figure 6C). Although the deletion of the PEST is expected to suppress the degradation of the AR, and hence induce higher protein levels, the lower activity of the deletion mutant is not surprising in view of the papers which describe the link between degradation and transcription control in case of the ER

(Reid et al. 2003) as well as the AR (see also further). The protein levels evaluated in Western blot do not seem to be affected by the deletion of the PEST sequence. An unexpected (small) decrease in signal in COS cells will have to be confirmed in additional assays.
The mutation of the Serine at position 650 into an Alanine prevents phosphorylation at this site (Zhou et al. 199; Wong et al. 2004)), mutation into a glutamate would mimic phosphorylation. In our hands, these changes did not affect the AR activity on the MMTV or the C3(1) based reporter genes in COS 7 cells or HeLa cells (Figure 6C and data not shown). This is in contrast with an earlier report (Zhou *et al.* 1995) but in agreement with an more recent screening for phosphorylation sites in the AR (Wong *et al.* 2004).
In conclusion, a link between the PEST-mediated degradation or the phosphorylation of Serine 650 and the superactivation of the AR resulting from the Δ1 or Δ11 deletion is unlikely.

II.E. Role of the hinge region in intracellular localization of the AR

To study the intracellular localization of the AR and mutants of it, we have fused the AR to the enhanced green fluorescent protein (EGFP). The first fusion protein had an impaired activity (see report of 2003). We obtained an EGFP fusion from Dr. Karen Knudsen (Ohio University, Cincinatti) in which a Gly-Ala linker separates the two proteins. In our hands, this AR-EGFP protein indeed has transactiviting properties which are similar to those of the AR (Figure 7A). When transfected cells are treated with 10 nM R1881, the AR fusion migrates from the cytoplasm to the nucleus (Figure 7B). However, the Δ1 construct seems to be excluded from the nucleus in the absence of hormone and redistributed upon hormone addition. Two types of cells can clearly be distinguished: cells with high intensity spots and cells with a more even distributed fluorescense signal. Surprisingly, in the presence of ligand, spots can be observed in the nucleus, while in the absence of hormone spots are apparent in the cytoplasm of some but not all cells (Figure 7B). It is unclear at this moment whether the spots are artefacts of the overexpression, and if not what the relevance and nature of these structures is in view of the higher activity of the Δ1 construct versus the wt AR.
We plan to monitor the localization of the AR during longer time periods to establish whether the different types of distribution are consecutive in time or not.

Conclusion: The Δ1 AR has an altered cellular distribution. This was expected, since part of the nuclear localization signal is deleted. It is surprising that this results in a more active AR.

II.F. Role of the hinge region in AR stability

In an earlier study, we analysed the effect of MG132, an inhibitor of the proteasome, on the activity of the AR versus the Δ1 construct (Tanner et al. 2004).

We have now followed the steady state of the AR versus Δ1 and Δ11 by Western blotting for AR in cellular extracts made 1, 6 and 24 hours after addition of hormone (Figure 8A). Although after one hour hormone stimulation, the effect is not outspoken, after six hours, the Δ1 and Δ11 constructs are clearly expressed to a higher level as compared to the wtAR. After 24 hours, higher bands become visible on the Western blot. They can not be explained by sumoylation (Callewaert et al. 2004). Possibly, these are explained by poly-ubiquitinated AR. This should be verified in immunoprecipitation assays with anti-AR antibodies, and analysis of the precipitate with anti-ubiquitine antibodies. Even in the absence of ligand, the Δ1 and Δ11 constructs are more expressed. This is not due to a enhancer-like element located in the hinge region coding cDNA part, since we did not observe androgen-regulation of a reporter gene with this fragment cloned upstream of the promotor (data not shown).

We have analysed in time the response to androgens in cells transfected with a C3(1)-based-luc reporter gene co-transfected with either wtAR, Δ1 or Δ11. The difference in responses to ligand on the C3(1) reporter are in parallel with the increase in receptor levels as observed in the Western blot (Figure 8B), except that, although Δ1 and Δ11 seem equally expressed, Δ11 is more active. These data need to be reproduced.

Conclusion: Expression levels of AR mutants can vary considerably. There is a close, but not perfect correlation between expression level and transactivation of the different deletion constructs. For mutant receptors with highest expression levels, poly-ubiquitinated forms become detectable.

II.G. Effect of hinge region on AF1

The results of section II.F forced us to re-evaluate the effects of the hinge on transactivation. Because the AF1 in the AR is strong and constitutive, meaning it can activate transcription in the

absence of the LBD, we focussed in this receptor fragment. The AF1 comprises two activation functions Tau 1 (between residues 100 and 360) and Tau 5 (between residues 370 and 485). We assayed the effect of the deletion of the hinge region on the activity of the NTD (Figure 9B). When we compare the constructs, depicted in figure 9A, truncated at 4 residues of the hinge region, the construct with the complete NTD is only 1.5-fold stronger compared to that containing the isolated tau5. Truncation of the first 171 or 100 residues results in a 4-fold stronger activation.

The lower activity of the full NTD is not observed with the CTE15 and CTE45 constructs. The four-fold increase between the tau5 and the other constructs is observed for the CTE15 constructs but not for the CTE 45 constructs.

The differences in activity are not reflecting relative expression levels of the constructs as determined in Western blotting (Figure 9C). This is evident e.g. for the constructs NTD CTE4 (a) and NTD CTE15 (b): (a) is more expressed, but (b) is more active. The Tau5 CTE15 construct seems even undetectable by Western blotting (two independent experiments), but still active in transcription. The DNA sequence of all constructs has been verified, but the data of the Western blot need to be further confirmed. The DNA binding is also being verified.

Conclusion: The deletion of the hinge affects the expression level of the protein, even when the LBD is not present, but this is not strictly correlated with the differences in activity. The fact that some CTE4 constructs are equally active as compared to CTE15 or CTE45 constructs is clearly discordant from their expression level.

II.H. Lysines 630, 632 and 633

The AR can be acetylated *in vivo* in the hinge region at Lysines 630, 632 and 633 (Fu *et al.* 2000, 2002; Pestell 2003). Lysine 630 was proposed to be involved in transcriptional regulation through an enhanced recruitment of p300 and reduced affinities for N-CoR and Smad3 after its acetylation. The mutation of this lysine to glutamine or threonine (mimics acetylation), when expressed in DU145 cells promoted cell survival and growth of cancer cells in soft agar and nude mice (Fu *et al.* 2003). More recently, the group of Robson proposed a link between Mdm2 mediated degradation and deacetylation (Gaughan et al. 2005). In the earlier reports, we showed data on AR mutated in several residues of the hinge. In view of the observations described in section II.G, the effects on transcription need to be correlated with expression levels.

We started by analysing AR with mutated Lysines 630, 632 and 633 into Valines and Alanines. Functional analyses are currently being done. The here reported results in Figure 10 are therefore fragmentary and preliminary. In HeLa cells, a double mutant in which Lysine 632 and 633 are exchanged for Valines is five times more active than wild type, while Δ1 is 15 times more active (Figure 10A). Here too, slightly different expression levels are observed in a first Western blot (figure 10B). Interesting to note is that longer exposures of the blot, there are bands of higher mobility in the lanes containing extracts from cells expressing Δ1 and Δ11. Similarly to those observed in Figure 8A, these bands might be explained by poly-ubiquitinated receptor. This will be tested after immunoprecipitation (see higher).

Earlier, we mutated Leucines 631 and 634 and Lysine 633 into Prolines. This did not affect the AR activity on a MMTV-reporter. Protein expression levels of these and other mutants still have to be checked in Western blot. A new mutation of Lysine 632 into Proline results in a loss of the predicted α-helix. This mutation has, however, has a small effect on the activity of the AR: a two to threefold higher induction (Figure 10A), and a slight increase in protein level as seen on the Western blot (Figure 10B).

Conclusion: Although the activity of AR with isolated or combined mutations was in some cases higher compared to wild type, it was never as outspoken as the effect seen for Δ1 or Δ11. The helical nature of this part of the hinge does not seem to be involved.

III. General conclusions

This is a point per point check of the milestones. The numbering refers to that of the 'Statement of work' of the grant application

First year

i.a. Crystals: have been made (report 2004)

i.b. Translation of data: has been done (report 2004) and a second generation constructs is being made/analyzed (this report, section I.A)

i.c. Deletion of the hinge region: has been done (earlier reports and conclusion in figure 1)

i.d. Make a yeast expression vector: has been done (earlier reports).

Second year

ii.a. Make a library of mutations in the hinge: has been done partially (report 2004) and will be continued with an adated mutagenesis protocol.

ii.b. Bacterial expression vectors: have been done in part (report 2004) and will be done when more information on the hinge will be available

ii.c. Prokaryotic expression for crystallization: has been done and will be continued when more information is available from mutation analyses (report 2003)

Third year

iii.a. Finalise screening in yeast: was proven unfeasable (report 2004). An alternative use of yeast is being developed in collaboration with Dr. Ceralin (Laboratoire de Cancerologie Experimentale et de Radiobiologie-EA/ULP 3430, IRCAD, Strasbourg, France) (see this report section IIA)

iii.b. Pro- and eukaryote expression constructs for specific mutants: has been done and is ongoing. **This has led to most of the observations reported in section II of this report.**

iii.c. Screening for dominant negative peptides. A first attempt has been reported in report 2004. This approach will be continued based on new information of the ongoing hinge region studies.

Fourth year

We are excited that we will be able to develop the experiments further untill April 2006. This should enable the confirmation of the hypothesis that the hinge is a major player in the control of the expression level of the AR. The mechanisms involved will also be studied in more detail (ubiquitination, cellular localization, communication with the two activation functions).

Abbreviations

AF activation function

AR androgen receptor

ARE androgen response element

CTE carboxyterminal extension

DBD DNA-binding domain

DBD-LBD AR fragment lacking the amino-terminal domain

EGFP enhanced green fluorescent protein

FRET fluorescence Resonance Energy Transfer

GR glucocorticoid receptor

MMTV mouse mammary tumour virus

MR mineralocorticoid receptor

NTD-DBD AR fragment lacking the ligand-binding domain

PR progesterone receptor

SRC 1 steroid receptor co-activator 1

Figure legends

Figure 1 : Introduction

1A. Figure from Shaffer et al. 2004 (appendix 1) depicting the AR-DBD dimer of subunit A and B, bound to a direct repeat ARE (lower double helix). The amino- and carboxyterminal ends are indicated by N and C resp. The 'spacer' of three nucleotides that separates the two 5'-TGTTCT-3' hexamers is indicated.

B. Schematic representation of the amino acid sequence of the human AR hinge region (amino acids 625 to 669). Constructs Δ1 and Δ11 encode the full-length hAR deleted of amino acids 628 to 646 and 628 to 636, respectively. The brackets indicate the CTE (carboxyterminal extension (Schoenmakers et al. 1999)

C. Transient transfections were performed in HeLa cells plated in 96-well plates at a density of 10^4 cells/well. 100 ng of the GRE-TAT-driven luciferase reporter construct was co-transfected with 10 ng of receptor expression plasmid and 10 ng of a pCMV-β-galactosidase construct. Receptor expression plasmids used encoded either the full-length wild-type AR (wtAR) or the deletion mutants Δ1 and Δ11, described above. Cells were incubated in the absence or presence of 10 nM R1881 (a synthetic androgen) for 24 hours. Cells were then harvested and luciferase and β-galactosidase values were measured. Relative inductions are the ratio's of the luciferase values (means of at least three independent experimants performed in triplicate, and corrected for β-galactosidase expression levels) of extracts from stimulated and unstimulated cells. Error bars indicate the standard errors of the mean.

Figure 2 : Translation of the structural data

A. Top right inset: picture taken from Shaffer etal. 2004 zooming in on the dimerisation interface of the two AR-DBD subunits bound to a direct repeat element. Top left inset: Sequence of the D-boxes of AR (left) and GR (right) the arrows indicate the two differing residues which were exchanged. The histograms display the results of transient transfections with AR or GR-derived constructs in which one or both residues were exchanged. Transfections and analyses were performed as described in the legend of figure 1. The histograms (PB-ARE2 and slp-HRE2) show no change in selectivity of GR or AR-derived constructs. The two left figures (TAT-GRE and IR3) indicate the mutant receptors are all able to transactivate through classical AREs.

B. Band shift assays were performed with isolated DBDs from AR or GR, or from mutant AR or GR-DBDs: AR muts: all three mutants in AR DBD: Ser to Gly, Thr to Iso and the combination.

GR muts: the reciproke mutations in the GR DBD. The upper panel display the titration curves (constant amount of DNA, increasing concentration of DBD) for the TAT-GRE (a classic ARE). The lower panels show the same for the selective PB-ARE2. All experiments were performed in triplicate and the error bars indicate standard errors of the mean. Clearly, the mutations did not result in major shifts in affinity for the TAT-GRE or the PB-ARE2.

Figure 3 : Progesterone receptor action through selective AREs

A. Alignment of the hinge regions of the human AR (hAR), the human progesterone receptor (hPR), the human mineralocorticoid receptor (hMR), and the human glucocorticoid receptor (hGR). Highly conserved amino acids are indicated by capital letters and less conserved amino acids are indicated by small letters.

B. The C3(1) ARE based luciferase reporter (left panel), or the probasin ARE based luciferase reporter (right panel) were transfected in HeLa cells at 100 ng per well (96-well with 10000 cells per well), together with 1 ng expression vector for the indicated receptors. The luciferase assays were performed after stimulation of the cells with 10 nM of the corresponding ligands, as described in the legend of figure 1C. The mean of three independent experiments performed in triplicate, +/_ steandard error of the mean are depicted.

C. Ligand titration assays to determine the minimal concentration of hormone needed to observe a response. Three reporters were transfected in HeLa cells: slpHRE 4x, contains four copies of the slpHRE2, PB-ARE2 contains two copies of the PB-ARE2, scARE1.2 contains four copies of the scARE1.2 elements (constructs described in Verrijdt et al. 2000).

Figure 4 : Effect of the hinge region on the AF2 function and the SRC1 interaction

A. The effect of the hinge region on the AF2 function was monitored by transient transfection experiments using the C-terminal part of the human AR consisting of the DBD, H and LBD (DBD-H-LBD) and a corresponding mutant construct with the deleted hinge region between aa 628 and 646 was created (gal4DBD-DBD-Δ-LBD). Similar gal4DBD fusion constructs were made (gal4DBD-DBD-H-LBD and gal4DBD-DBD-Δ-LBD). Transient transfection experiments were performed in HeLa cells as described in the legend of figure 1: 100 ng of either the C3(1) ARE or the gal4 UAS luciferase reporter construct was co-transfected with 10 ng of gal4DBD plasmid and 10 ng of a pCMV-β-galactosidase construct. Cells were stimulated for 24 hours with 10 nM R1881. The expression of the AR fragments were confirmed by Western blotting. Equal expression was observed.

B. The effect of the hinge region on the SRC1 interaction was examined by double hybrid experiments using a SRC fragment (aa 1241-1441). 100 ng of the gal4 UAS luciferase reporter constuct was cotransfected with 10 ng of a pCMV-β-galactosidase construct, 50 ng of a gal4DBD SRC1a fusion construct and 50 ng of either the wild type (DBD-H-LBD-VP16) or the mutant AR-VP16 fusion plasmid (DBD-Δ-LBD-VP16). Cells were stimulated and harvested as in A. Luciferase values given are relative to the activity of the gal4DBD co-expressed with VP16.

Figure 5 : Effect of the hinge region on the N/C interaction A.The effect of the hinge region on the N/C interaction was evaluated by comparing the transcriptional activity of the wild type (WT) and mutant full size AR (Δ1) in transient transfection experiments with the activity of corresponding constructs with a deleted FQNLF motif in the NTD (ΔFQNLF and (ΔFQNLF-Δ1): COS cells were transfected with 100 ng of a MMTV luciferase reporter, 10 ng of a pCMV-β-galactosidase construct and 10 ng of a full size AR expression plasmid. Cells were stimulated and harvested as in figure 3A.
B.The effect of the hinge region on the N/C interaction was examined by double hybrid experiments in COS cells: 100 ng of the gal4 UAS luciferase reporter constuct was co-transfected with 10 ng of a pCMV-β-galactosidase construct, 50 ng of either a wild type (gal4DBD-DBD-H-LBD) or a mutant gal4DBD fusion construct (gal4DBD-DBD-Δ-LBD) and 50 ng of either the wild type NTD (NTD-VP16) or the mutant NTDVP16 fusion plasmid (ΔFQNLF-VP16). Cells were stimulated and harvested as in figure A. The luciferase activity are represented as absolute luciferase values.

Figure 6 : A putative PEST sequence A. Schematic representation of the full size and ΔPEST AR. The amino acid sequence of the wild type hinge region (aa 628-669) is shown. The PEST domain situated between aa 638 and 658 is deleted in the ΔPEST construct.
B. The effect of the PEST sequence on the transcriptional activity was evaluated by comparing the activity of the WT and mutant ΔPEST construct in transient transfection experiments. 10^4 HeLa or COS cells plated in 96-well plates, were transfected with 100 ng of the indicated luciferase reporter 10 ng of a pCMV-β-galactosidase construct and 10 ng of a full size AR expression plasmid. Cells were stimulated and harvested as in figure 3A. The expression of the constructs was confirmed by Western blotting (inset at the right).
C. Similar to section B, the effect of mutations in Lysine 638 and 658 as well as in Serine 650 was assayed on an MMTV-luciferase reporter in COS and HeLa cells.

Figure 7 : Effect of the hinge region on the intracellular localisation A. Schematic presentation and transcriptional activity of the C-terminal fusion of both WT and Δ1 constructs with EGFP containing a linker consisting of a Gly-Ala repeat.
The transcriptional activity of EGFP-WT and EGFP-Δ1 was evaluated and compared with the activity of the flag-tagged WT and Δ1 constructs in transient transfection experiments. 10^4 HeLa or COS cells plated in 96-well plates, were transfected with 100 ng of the MMTV luciferase reporter, 10 ng of a pCMV-β-galactosidase construct and 10 ng of an AR expression plasmid. Analysis of data was done as described in the legend of figure 1.
B. Fluorescence microscopy of the EGFP-WT and -Δ1 constructs. 10^4 HeLa cells were transienty transfected with 500 ng of the indicated construct and stimulated for 1 h hour with 10 nM R1881 when mentioned.

Figure 8. The effect of the hinge region on the stability of the AR A. The constructs wtAR, Δ1 and Δ11 (described in figure 1) were expressed in HeLa cells seeded into 24-well plates at 10^5 cells/well. Twenty-four hours after transfection, cells were incubated in the absence or presence of 10 nM R1881. Extracts were prepared either 1, 6 or 24 hours after addition of hormone. Equal amounts of extracts were loaded onto a 10% SDS-polyacrylamide gel, electrophoresed and blotted onto Hybond-P nitrocellulose membrane. All the full-length constructs are tagged at the N-terminus with the Flag peptide and so the membrane was probed with anti-Flag antibody, receptor proteins detected by ECL and autoradiography performed. The arrowhead indicates possible poly-ubiquitinated AR.
B. Transient transfections were performed in HeLa cells as described in figure 1. Here 100 ng of the C3(1)-based luciferase reporter construct was co-transfected with either 10 ng of wtAR, Δ1 or Δ11 expression plasmid and 10 ng of a pCMV-β-galactosidase construct. Cells were incubated in the absence or presence of 10 nM R1881 for the indicated time intervals before being harvested for assays. The inserted panel is a magnification of relative inductions obtained at 3 and 6 hours, demonstrating a response to ligand that parallels the increase in expression observed after 6 hours.

Figure 9. The effect of the hinge region on AF1 A. Schematic representation of the amino acid sequence of the human AR hinge region (amino acids 625 to 669). Constructs Δ1 and Δ11

as described in figure 1. Constructs described with CTE4, CTE15 and CTE45 are truncated at amino acids 628, 639 and 669 respectively.

B. Transient transfections were performed in HeLa cells as described in figure 1. Here 100 ng of the GRE-TAT-luciferase reporter construct was co-transfected with 10 ng of the various NTD-CTE expression plasmids and 10 ng of a pCMV-β-galactosidase construct. The constructs contain one of the three possible C-terminal truncations (CTE4, CTE15 or CTE45) together with one of the possible N-terminal truncations. With regards to the N-terminus, constructs express either: (i) the full-length N-terminal domain (NTD); (ii) Tau 5 starting at amino acid 360 (360-NTD); (iii) truncated Tau 1 starting at amino acid 171 (171-NTD); or (iv) the full Tau 1 starting at amino acid 100 (100-NTD). Cells were incubated in the absence of hormone (as all constructs are devoid of the ligand binding domain) for 24 hours before being harvested for assays. Data represent the means of at least three independent experiments performed in triplicate. Luciferase values are corrected for β-galactosidase expression levels and expressed as relative light units (rlu).

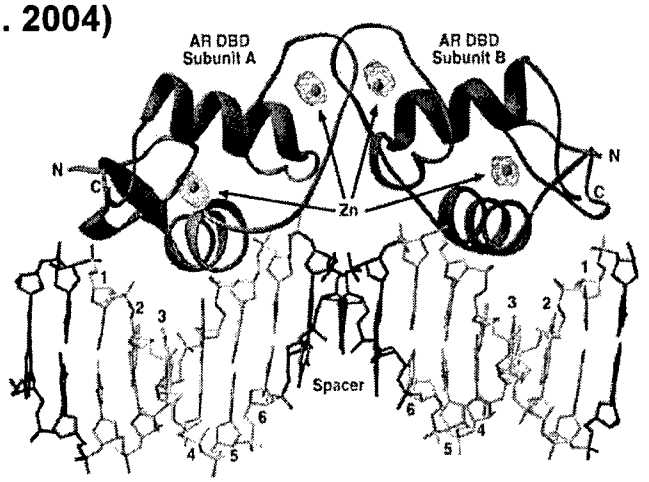
C. The NTD-CTE constructs (tagged at the N-terminus with the Flag peptide) were expressed in HeLa cells seeded into 6-well plates at $3x10^5$ cells/well. Extracts were prepared after a 24 hour incubation in the absence of hormone and expression analysed by Western Blotting as described in the legend of figure 8A.

Figure 10. The role of Lysine 632 and 633 A. Transient transfections were performed in HeLa cells as described in figure 1: 100 ng of the GRE-TAT-luciferase reporter construct was co-transfected with 10 ng of the various full-length AR expression plasmids and 10 ng of a pCMV-β-galactosidase construct. The constructs used were : (i) full-length wild-type AR (wtAR); (ii) AR deleted of amino acids 628-646 (Δ1); (iii) AR with lysine 632 mutated to a valine (K632V); (iv) AR with both lysines 632 and 633 mutated to valine (K632/K633V); and (v) AR with lysine 632 mutated to a proline (K632P). Cells were incubated in the absence and presence of 10 nM R1881. The relative inductions represent the means of at least three independent experiments performed in triplicate.

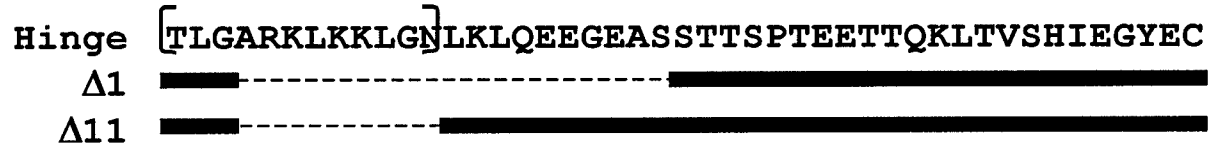
B. The constructs encoding wtAR, Δ1, Δ11, K632V, K632/K633V, and K632P (tagged at the N-terminus with the Flag peptide) were expressed in HeLa cells seeded into 10 cm plates at $2x10^6$ cells/plate. Extracts were prepared after a 24 hour incubation in the presence of 10 nM R1881 and expression analysed by Western Blotting as described in figure 8.

Figure 1 Introduction

A. Structural data on AR DBD binding to a direct repeat (Shaffer etal. 2004)



B. Hinge deletions used in this study



C. Functional comparison

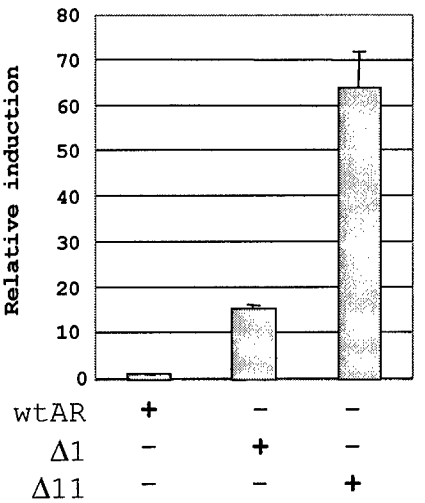
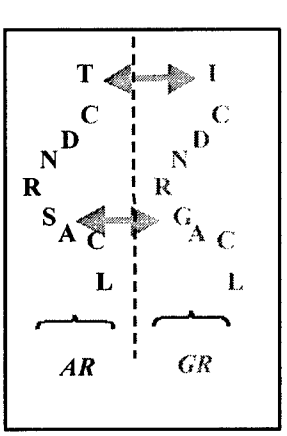
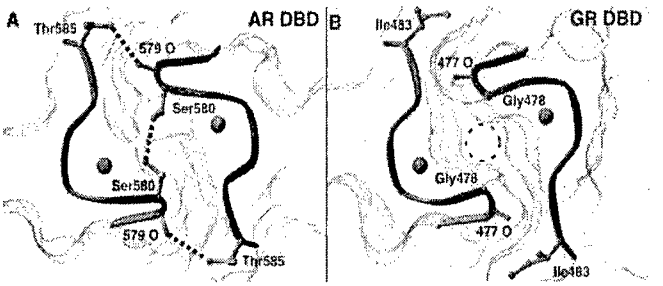
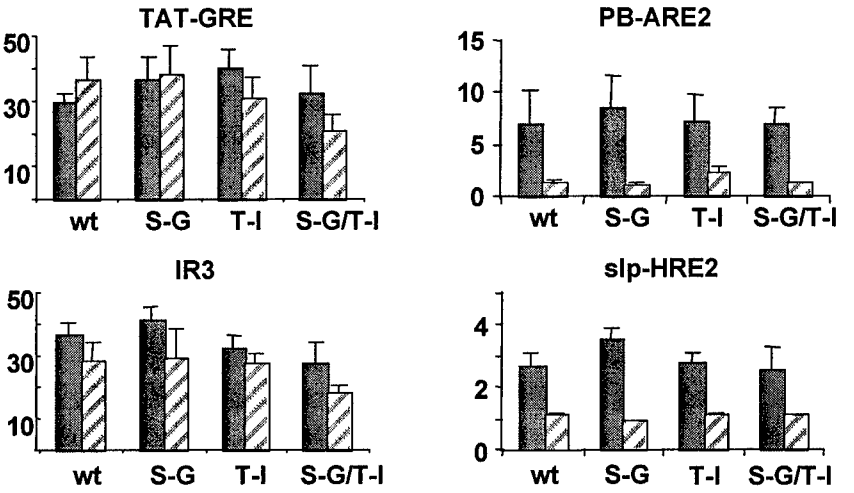


Figure 2 A. Functional analysis of AR and GR mutations in the D-box







AR GR

B. DNA binding characteristics of wil type and mutated DBDs

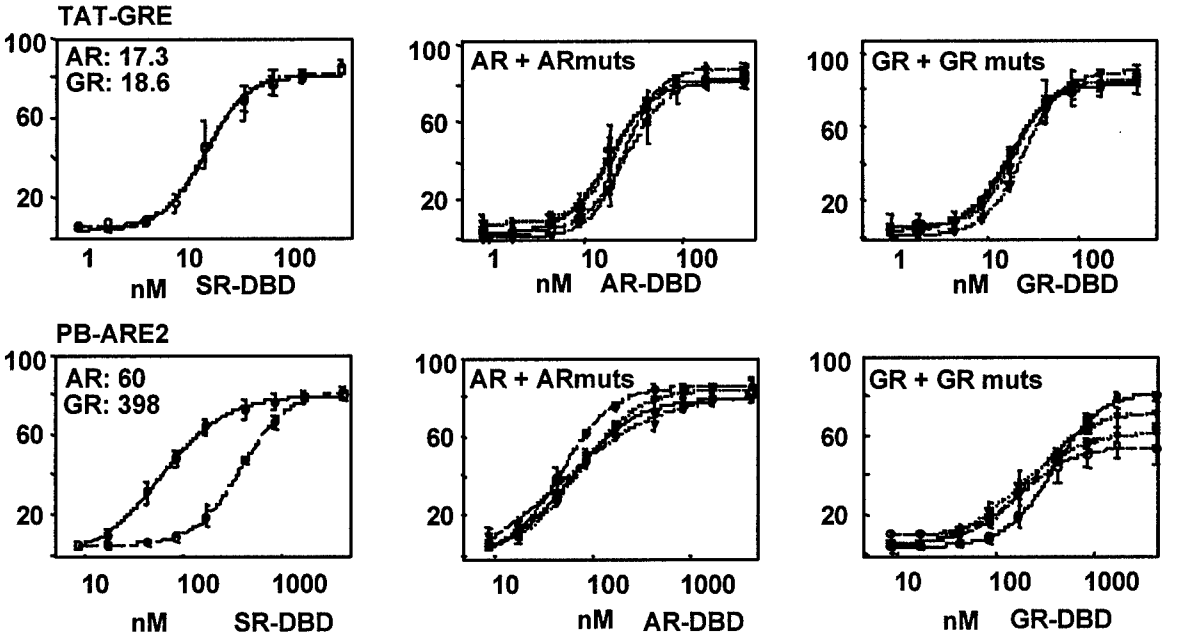
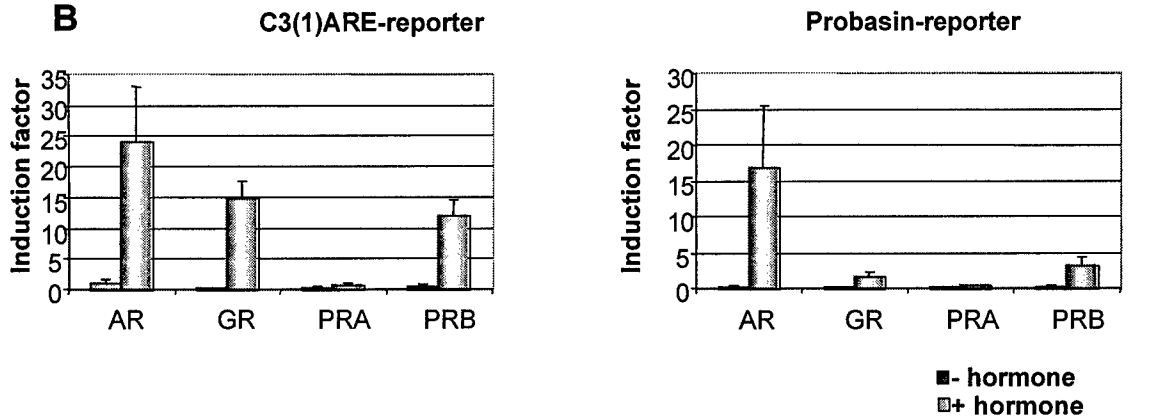


Figure 3: Progesterone receptor action through selective AREs

A

```
                1        10        20        30        40        50        60      66
                |--------+---------+---------+---------+---------+---------+-----|
hAR             TLGARKLKKLGNLKLQEEGEASSTTSPTEETTQKLTVSHIEGYEC
hPR             VLGGRKFKKFNKVRVVRALDAVALPQPLGVPNESQALSQRFTFSPGQDIQL
hMR             NLGARKSKKLGKLKGIHEEQPQQQQPPPPPPPPQSPEEGTTYIAPAKEPSVNTALVPQLSTISRAL
hGR             NLEARKTKK--KIKGIQQATTGVSQETSENPGNKTIVPATLPQL
Consensus       nLgaRK.KK..k.kgi........q.p.e.p..k..v..t..........................
```



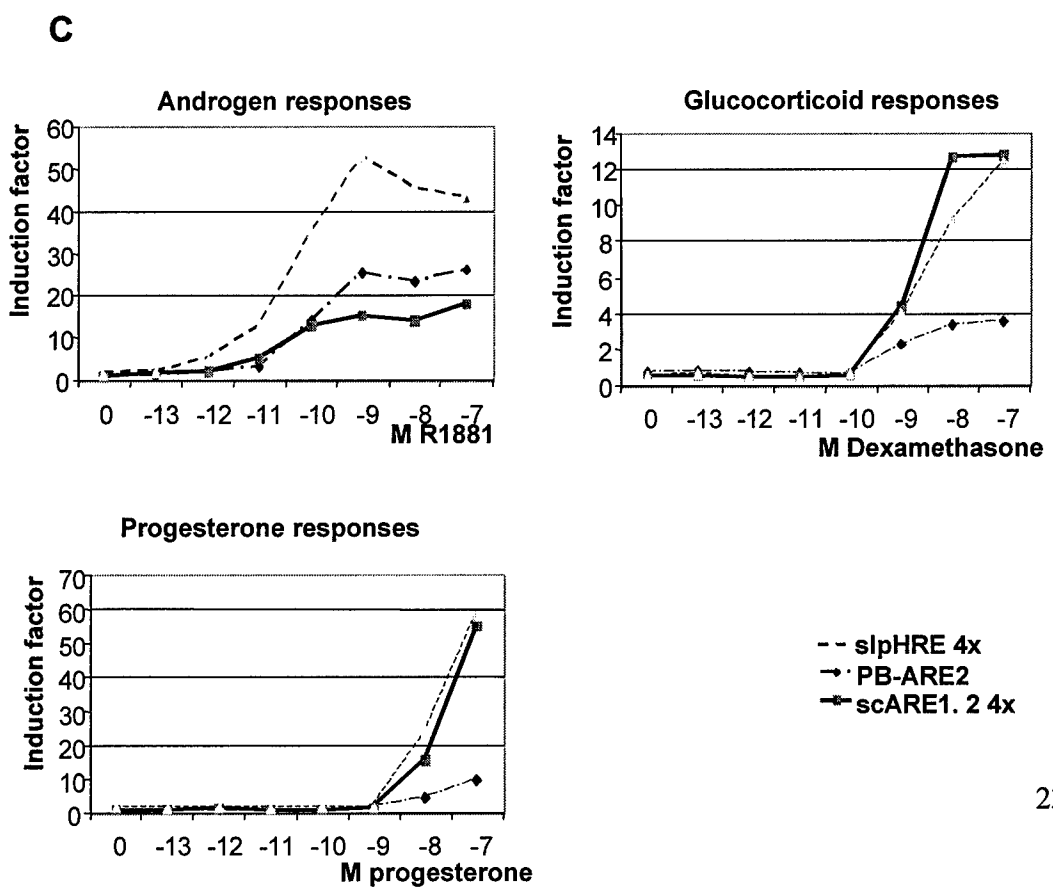
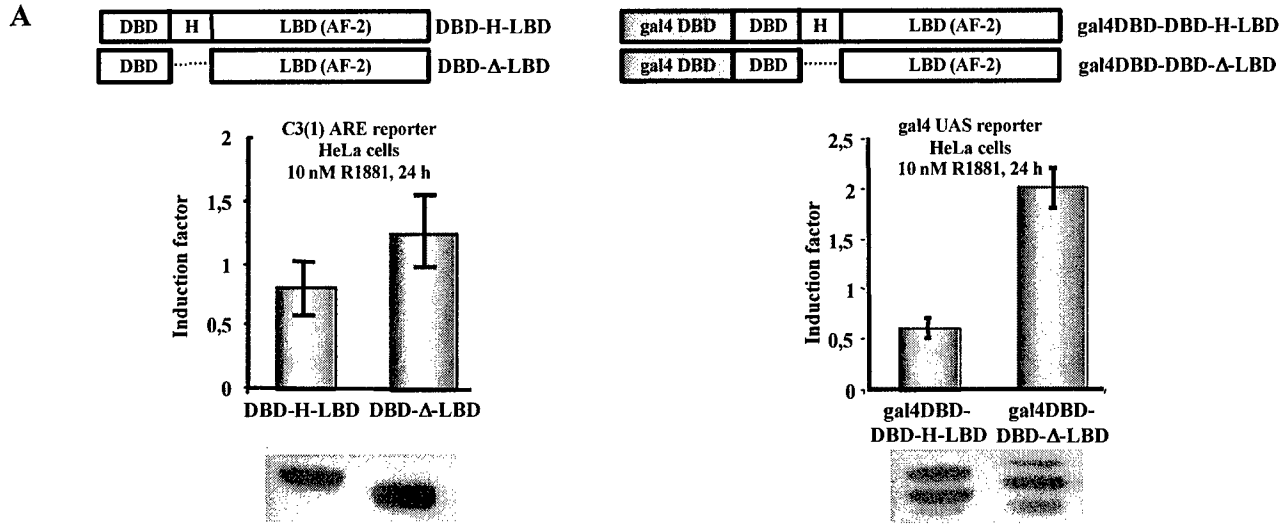


Figure 4 Effects of the hinge on the LBD



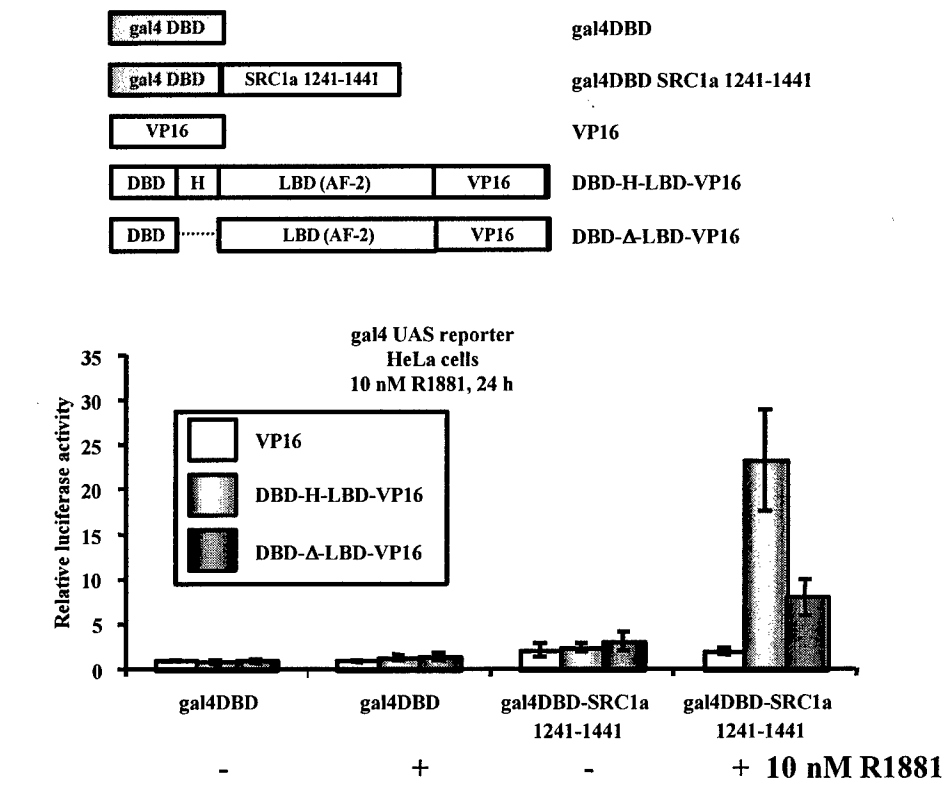


Figure 5: The hinge and the N/C interactions

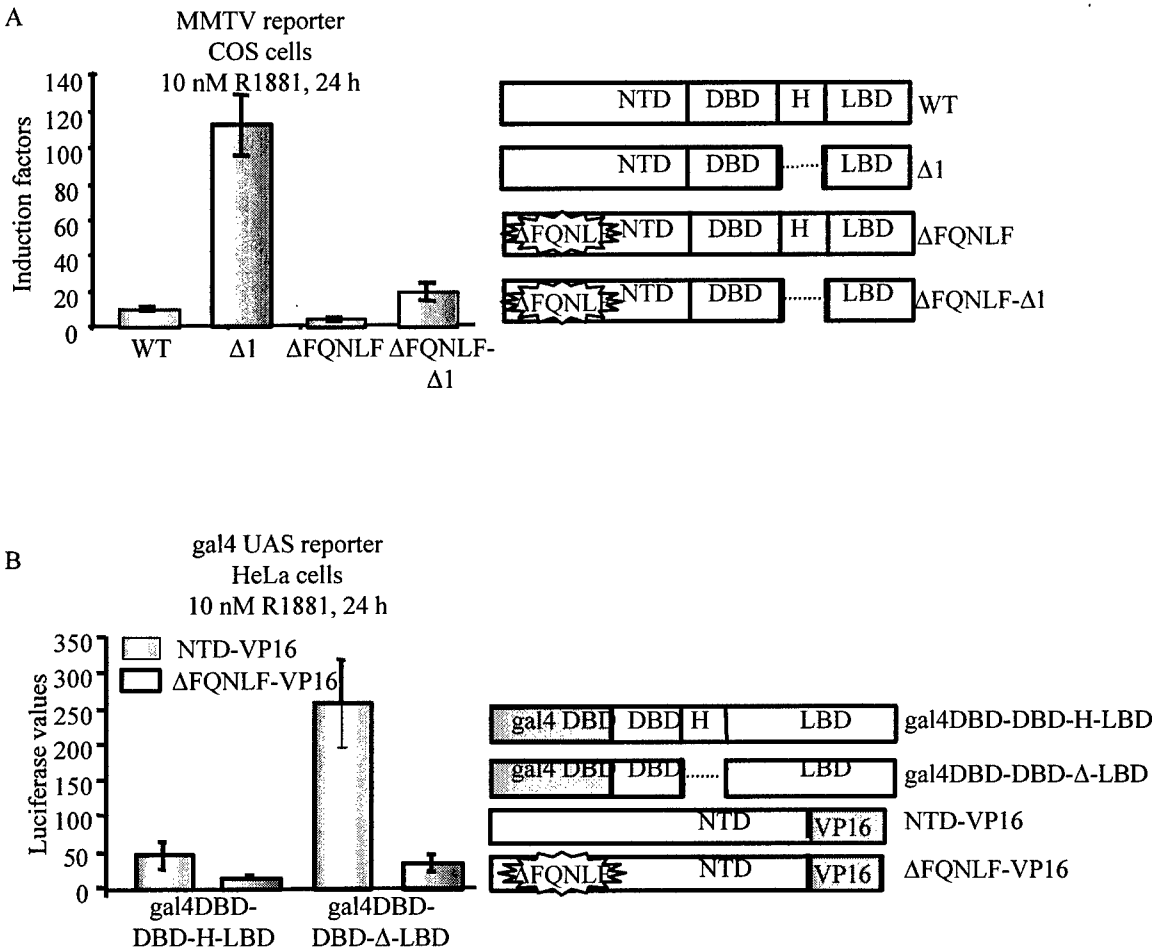


Figure 6: a putative PEST sequence in the hinge

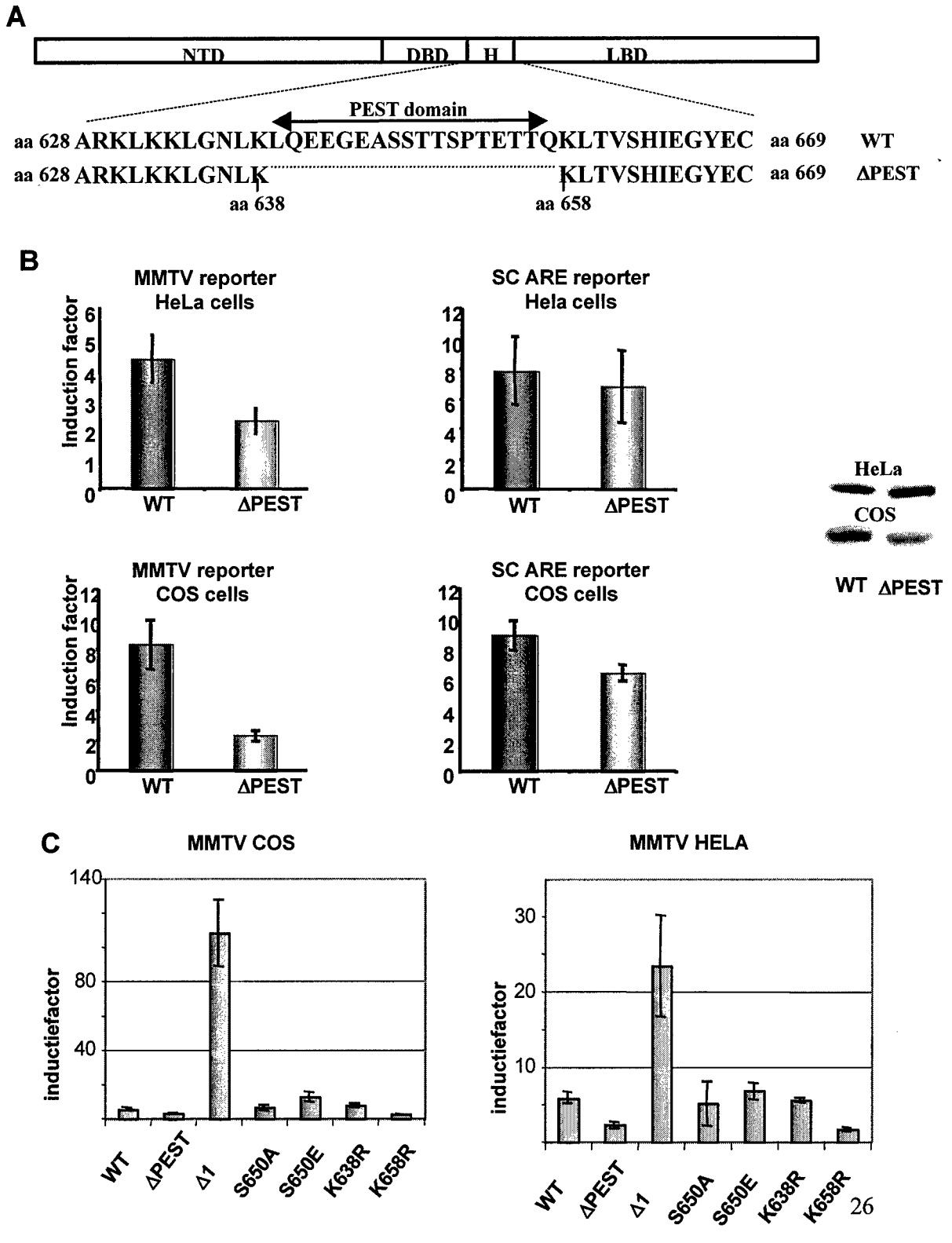
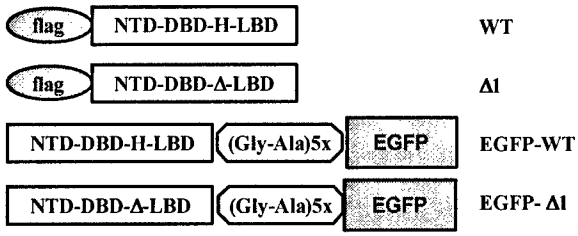
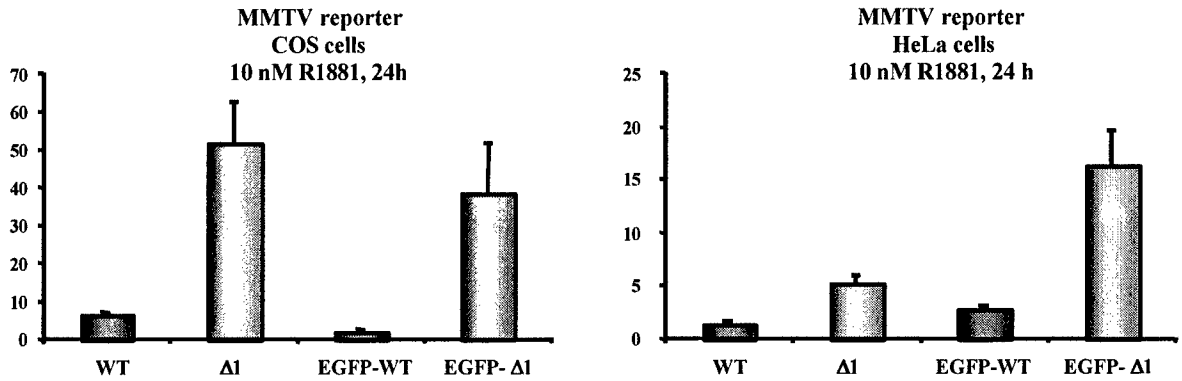


Figure 7 : Cellular localization of AR

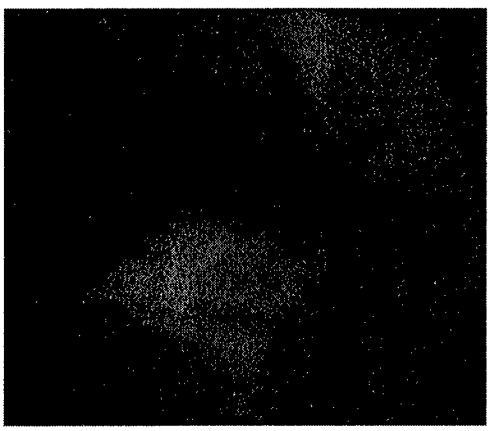
A



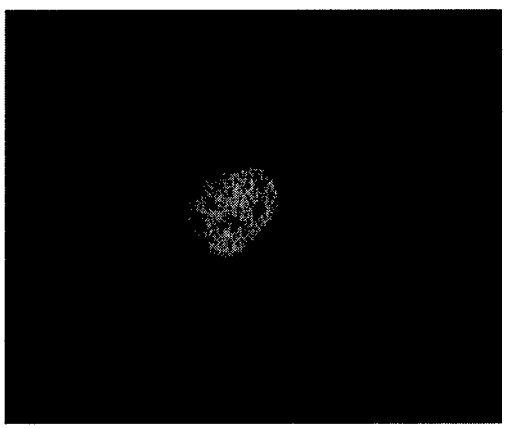


B

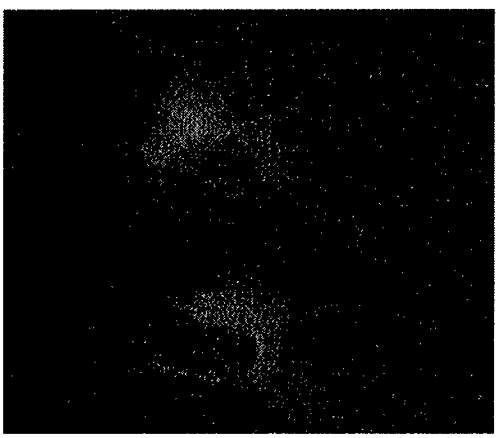
EGFP-WT, no stimulation



EGFP-WT, 10 nM R1881, 1 h



EGFP- Δ1, no stimulation



EGFP- Δ1, 10 nM R1881, 1 h

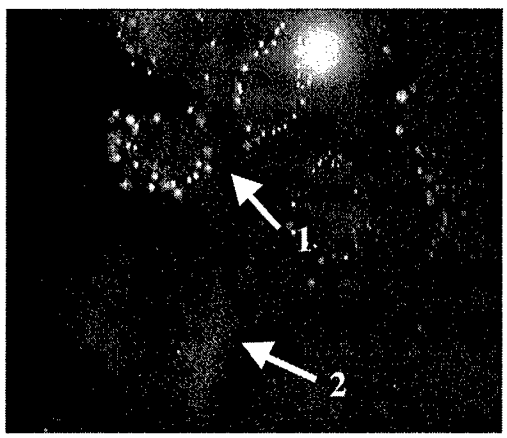
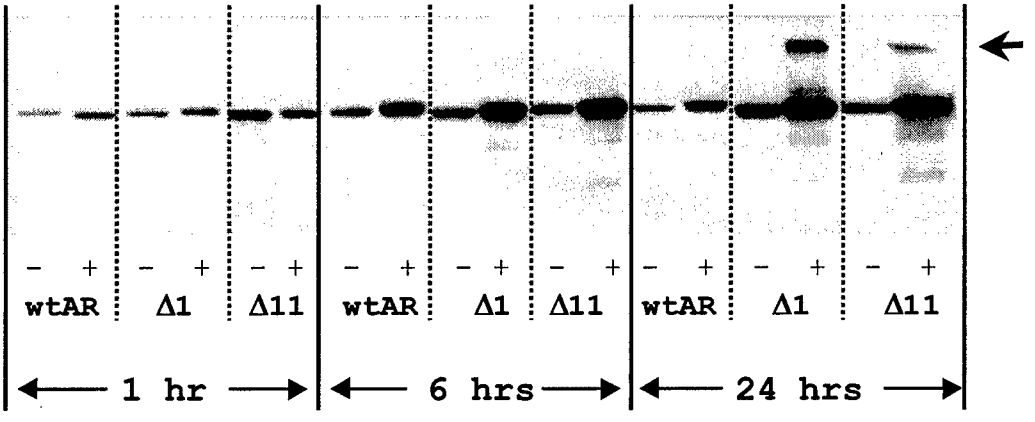


Figure 8

A. Western blot showing time course of receptor expression



B. Normalized luciferase values time course of receptor expression

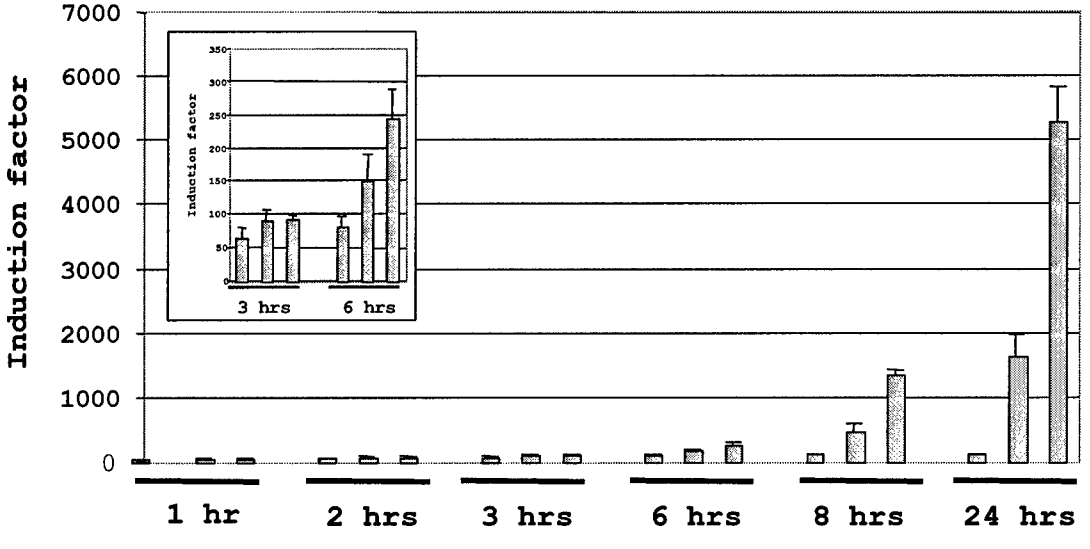
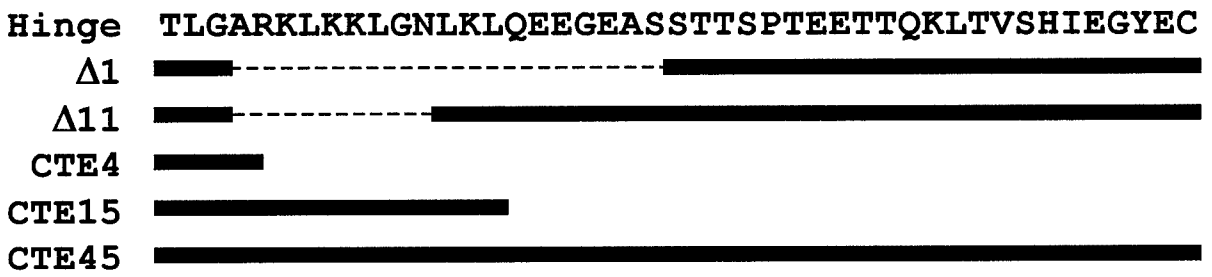
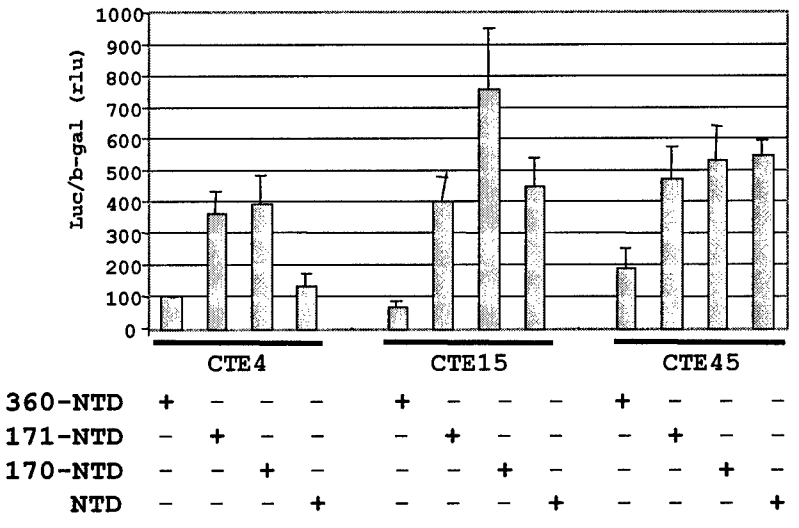


Figure 9

A. Deletion constructs used in panel B and C



B. Functional implications of varying the CTE



C. Western blot showing differences in expression levels

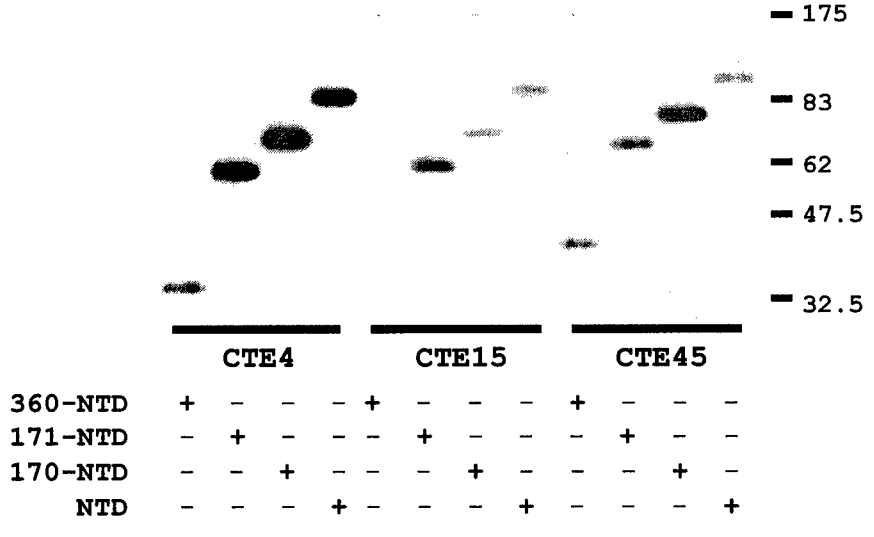
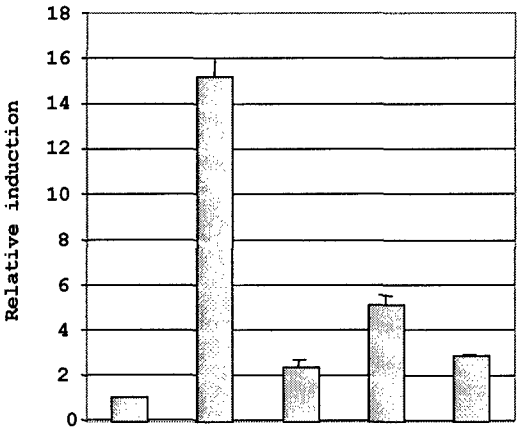


Figure 10 Correlation between androgen response and expression level

A



wtAR	+	–	–	–	–
Δ1	–	+	–	–	–
K632V	–	–	+	–	–
K632/K633V	–	–	–	+	–
K632P	–	–	–	–	+

B

wtAR	+	–	–	–	–	–
Δ1	–	+	–	–	–	–
Δ11	–	–	+	–	–	–
K632V	–	–	–	+	–	–
K632/K633V	–	–	–	–	+	–
K632P	–	–	–	–	–	+

Key Research Accomplishments

The hinge region is a key control element of the androgen receptor. An eight residue long fragment carboxyterminal of the second zinc finger is involved in the recognition of selective androgen response elements.
We have established that the PR can also activate transcription through the selective AREs, albeit that the concentration of progesterone in male serum is lower than the lowest dose needed for activation through slective AREs.
The eight amino acid motif has a role in the attenuation of both activation functions of the AR (AF1 and AF2), in nuclear localization and finally in the control of expression level of the AR.
A PEST motif, located carboxyterminal of the eight amino acid motif has no major role in the control of AR degradation, but a deletion of it represses the androgen response mildly.

We have optimized a protocol which should allow easier introduction of mutations in the hinge region and transfer of mutations into yeast expression vectors.
We have also obtained a fusion between AR and EGFP which is as active as the AR in functional assays and in which the deletion of the eight amino acid motif has an effect comparable to that in wild type AR. This construct will now be used to follow the AR in time during androgen induced transcription and cellular movements.

We are obliged to the organization that we are allowed to continue our research untill April 2006. This extension will allow us to continue our research and should be sufficient to meet with the milestones put forward in the 'Statement of work'.

Reportable outcomes:

Papers in International refereed journals

Shaffer, P.L., Jivan, A., Dollins, D.E., Claessens F., Gewirth P. Structural basis of androgen receptor binding to selective androgen response elements. . *P. N.A.S. USA*, 101, 4758-4763, 2004,

Tanner, T., Claessens, F., Haelens, A. (2004) The hinge region of the androgen receptor plays a role in proteasome-mediated transcriptional activation. Ann. N.Y.Acad. Sci. 1030, 586-590,.

Christiaens, V., Berckmans P., Haelens A., Witters H. and Claessens, F. Comparison of different androgen bioassays in the screening for environmental (anti)androgenic activity. Environmental Toxicology and Chemistry accepted for publication on 20/04/2005

Lectures

Verrijdt, G., Peeters, A., Schauwaers, K. and Claessens, F. (2004) Mutational analysis of the dimerisation interfaces of the androgen and glucocorticoid receptor DNA-binding domains. Bioscience 2004 Meeting 'From molecules to organisms' Glasgow, UK, july 18 – 22th 2004.

Claessens F. Mécanismes moléculaires de l'action des androgènes. Journée d'Endocrinologie Sexuelle Alfred Jost, 7 februari 2004, Hôpital Cochin, Parijs

Claessens F. Molecular basis of androgen selectivity. Organon Oss Netherlands 22 September 2004.

Gewirth D. Structural basis of androgen receptor binding to selective androgen response elements. International Androgen 2004 Symposium at Berlin, 8 October 2004.

Claessens F. A crystal clear message on selective androgen response elements. International Androgen 2004 Symposium at Berlin, 8 October 2004.

Claessens F. The molecular biology of the androgen receptor: NTD, DBD, LBD and most of all 'the hinge' Lecture at the CelGen Division of Medical Faculty of the KULeuven.

Poster presentations at international meetings

Callewaert, L., Verrijdt, G., Haelens, A., Claessens, F. Different action mechanisms of the androgen receptor on selective versus canonical androgen response elements. Nuclear receptors. Stokholm, Zweden, 10-13 oktober 2004.

Callewaert, L., Verrijdt, G., Haelens, A., Claessens, F. Different action mechanisms of the androgen receptor on selective versus canonical androgen response elements. Androgens 2004. Symposium on androgen receptor function. Berlijn, Duitsland, 7-8 oktober 2004.

Internship reports

Master thesis in Biomedical Sciences Functionele analyse van de hinge-regio van de androgeenreceptor. Eindwerk voorgedragen tot het behalen van de graad van licenciaat in de Biomedische Wetenschappen door Kelly Gijsemans. (academiejaar 2003-2004)

Master thesis for Industrial Ingenieur Studie van de hinge-regio op de transcriptie-activatie van de humane androgeenreceptor. Ondernemingsproject voorgedragen tot het behalen van de graaad van industrieel ingenieur door Kelly Van der Sande. (academiejaar 2003-2004)

Master thesis in Pharmaceutical Sciences Annelies Peeters "De androgeen receptor: de invloed van D-box mutaties op DNA-herkenning"
Eindverhandeling ingediend tot het behalen van het Diploma van Apotheker.
Katholieke Universiteit Leuven, Faculteit Farmaceutische Wetenschappen.academiejaar 2003-2004

PhD thesis

PhD in Medical Sciences Leen Callewaert "Structure/function analysis of the amino-terminal domain of the androgen receptor"
Public defens and date of the degree 29 March 2004

PhD in Pharmaceutical Sciences Valerie Christiaens "Modulation of androgen receptor activity by p160 coactivators and a study of environmental contaminants"
Public defens and date of degree 23 March 2005

Conclusions

1. **The androgen receptor DNA binding domain co-crystallizes with a direct repeat of the 5'-TGTTCT-3' hexamer separated by three nucleotides in a head-to-head conformation**. One monomer binds with high affinity to a 5'-TGTTCT-3' hexamer, while the other recognises the complementary strand in the second hexamer (5'-AGAACA-3'), in which the position of G and C are identical to those in the 5'-TGTTCT-3'-sequence. The lower affinity binding seems to be compensated for by a stronger dimerization interface in case of the AR, not in case of the GR.

Mutation analyses have failed to confirm this hypothesis, and we will now return our attention to the carboxy-terminal extension (CTE).

2. There is a higher similarity between AR and PR in the CTE than between GR and AR or GR and PR. We therefore tested whether **PR is also able to act through selective AREs**. Although it clearly can, we still prefer to call the elements selective AREs, since the hormone concentration needed to activate the PR is higher than that normally observed in male serum.

3. **A deletion of the hinge region results in an AR which is more potent**. This is due to the deletion of eight amino acids (ARKLKKLGN). This deletion affects both activation functions AF1 and AF2, as well as the interaction between them.

4. **A putative PEST sequence** is located carboxy-terminal of the eight amino acid motif. The deletion of this PEST sequence results in a less active AR when tested on different reporter constructs. A mutation of Lysine 658 into arginine also attenuates the AR activity. This contrasts with the superactivity of the eight amino acid deleted AR. Maybe this observation is linked with the fact that degradation of nuclear receptors is linked with their transcription activation.

5. Since the eight amino acid motif overlaps with the nuclear localization signal, we have done **localization studies with EGFP-fused ARs**. This should enable us to follow in time the cellular distribution of AR and its mutant forms. A Gly-Ala linker between the two proteins is necessary, since a direct fusion protein was largely impaired in its activity (not shown).

6. **The deletion of the eight amino acids has a dramatic effect on the steady state of the androgen receptor**. The deletion results in a build up of higher protein levels which becomes apparent 6 hours after stimulation with hormone. In transactivation assays, the higher activity of the AR without the eight amino acids parallels the higher concentrations.

7. The effect described in 6 is also apparent when the NTD-DBD is tested. A carboxyterminal extension of 4 residues results in a higher expression compared with constructs having 15 or 45

residue long CTEs. **Surprisingly, the activity of the constructs does not seem to parallel exactly the differences in protein levels**. These expression studies will have to be confirmed.
8. Mutation analysis of the eight amino acids indicates **a role for lysine 632 and Lysine 633**, but the mechanism of action is still unclear. Changing Lysine 632 into Proline had no larger effect as compared to mutation into Valine. Mutations into other residues mimicking or preventing acetylation/methylation and combinations of mutations are planned in view of recent findings linking phosphorylation in the aminoterminal domain with deactelyation of the AR hinge.

References

Alen, P., Claessens, F., Verhoeven, G., Rombauts, W. and Peeters, B. (1999) The androgen receptor amino-terminal domain plays a key role in p160 coactivator-stimulated gene transcription. *Mol. Cell. Biol.* **19**: 6085-6097

Barbulescu, K., Geserick, C., Schuttke, I., Schleuning, W.D. and Haendler B. (2001) New androgen response elements in the murine pem promotor mediate selective transactivation *Mol. Endocrinol.* **15**, 1803-1816.

Bevan, C., Hoare, S., Claessens, F., Heery, D.M., Parker, M.G. (1999) The AF1 and AF2 domains of the androgen receptor interact with distinct regions of SRC1. *Mol. Cell. Biol.* **19**: 8383-8392

Callewaert, L., Verrijdt, G., Haelens, A., Claessens, F. (2004) Differential effect of sumoylation of the androgen receptor in the control of cooperativity on selective versus canonical response elements. *Mol. Endocrinol.* **18**: 1438-1449

Ceraline J, Erdmann E, Erbs P, Deslandres-Cruchant M, Jacqmin D, Duclos B, Klein-Soyer C, Dufour P, Bergerat JP. (2003) A yeast-based functional assay for the detection of the mutant androgen receptor in prostate cancer. *Eur J Endocrinol.* **148**: 99-110.

Christiaens, V., Bevan C., Callewaert, L., Haelens, A., Verrijdt, G., Rombauts, W., Claessens, F. (2002) Characterization of the two coactivator interacting surfaces of the androgen receptor and their relative role in transcriptional control. *J. Biol. Chem.* **277**, 49230-37,

Claessens, F., Alen, P., Devos, A., Peeters, B., Verhoeven, G. and Rombauts, W. (1996) The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors. *J. Biol. Chem.* **271**, 19013-19016

Claessens, F., Gewirth, D. (2004) The nuclear receptor superfamily. *Essays in Biochemistry* **Volume 40** 59-72.

Claessens F, Verrijdt G, Schoenmakers E, Haelens A, Peeters B, Verhoeven G, Rombauts W (2001) Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation. *J. Steroid. Biochem. Mol. Biol.* **76**:23-30.

De Gendt, K., Swinnen, J.V., Saunders, P.T.K., Schoonjans, L., Dewerchin, M., Devos, A., Tan, K., Atanassova, N., Claessens, F., Lécureuil, C., Heyns, W., Carmeliet, P., Guillou, F., Sharpe, R.M., Verhoeven, G. (2004) A sertoli cell-selective knock out of the androgen receptor causes spermatogenic arrest in meiosis. *Proc. Natl. Acad. Sci U.S.A.* **101**: 1327-1332.

Dubbink, H.J., Hersmus, R., Verma, C.S., van der Korput, H.A., Berrevoets, C.A., van Tol, J., Ziel-van der Made A.C., Brinkmann A.O., Pike A.C., Trapman, J. (2004) *Mol. Endocrinol.* **18,** 2132-2150.

Fu, M., Rao, M., Wang, C., Sakamaki, T., Wang, J., Di Vizio D., Zhang, X., Albanese, C., Balk, S., Chang, C., Fan, S., Rosen, E., Palvimo, J., Janne, O.A., Muratoglu, S., Avantaggiati, M.L.,

Fu, M., Wang, C., Reutens, A.T., Wang, J., Angeletti, R.H., Siconolfi-Baez, L., Ogryzko, V., Avantagiatti, M.L., Pestell R.G. (2000) p300 and p300/cAMP-response element-binding protein-associated factor acetylate the androgen receptor at sites governing hormone-dependent transactivation. *J. Biol. Chem.* **275**: 20853-860.

Fu, M., Wang, C., Wang, J., Zhang, X., Sakamaki, T., Yueng, Y.G., Chang, C., Hopp, T., Fuqua, S.A.W., Jaffray, E., Hay, R.T., Palvimo, J., Janne, O.A., Pestell, R.G. (2002) Androgen receptor acetylation governs trans activation and MEKK1-induced apoptosis without affecting in vitro sumoylation and trans-repression function *Mol. Cell. Biol.* **22**:3373-3388.

Gaughan L., Logan, I.R., Neal D.E., Robson, C.N. (2005) Regulation of androgen receptor and histone deacetylase 1 by Mdm2-mediated ubiquitylation. *Nucl. Acids Res.* **33**, 13-26

He, B., Wilson E.M. (2002) Mol. Genet. Metab. 75, 293-298.

Jenster G., van der Korput, H.A., Trapman, J., Brinkmann A.O. (1995) Identification of two transcription activation units in the N-terminal domain of the human androgen *receptor J. Biol. Chem.* **270**, 7341-7346.

Moilanen A, Rouleau N, Ikonen T, Palvimo JJ, Janne OA (1997) The presence of a transcription activation function in the hormone-binding domain of androgen receptor is revealed by studies in yeast cells. *FEBS Lett.* **412**:355-358.

Pestell, R.G. (2003) Acetylation of androgen receptor enhances coactivator binding and promotes prostate cancer cell growth. *Mol. Cell. Biol.* **23**: 8563-8575.

Reid G, Hubner MR, Metivier R, Brand H, Denger S, Manu D, Beaudouin J, Ellenberg J, Gannon F (2003) Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling. *Mol. Cell.* **11**:695-707.

Schneider, JS, Stone, MK, Wynne-Edwards, KE, Horton, TH, Lydon, J, O'Malley, B, Levine, JE (2003) Progesterone receptors mediate male aggression towards infants *Proc. Natl. Acad. Sci U.S.A* **100**, 2951-2956.

Schoenmakers, E., Alen, P., Haelens, A., Verrijdt, G., Peeters, B., Rombauts, W. and Claessens, F. (1999) Identification of features within the DNA-binding domain of the androgen and glucocorticoid receptor responsible for the discrimination between response elements. *Biochem. J.* **341**: 515-521

Schoenmakers, E., Verrijdt, G., Peeters, B., Verhoeven, G., Rombauts, W. and Claessens F. (2000) Differences in DNA binding characteristics of the androgen and glucocorticoid receptor can determine hormone specific responses. *J. Biol. Chem*. **275**, 12290-12297.

Shaffer, P., Jivan, A., Dollins, D.E., Claessens, F., Gewirth, D. (2004) Structural basis of androgen receptor binding to selective androgen response elements. *Proc. Natl. Acad. Sci U.S.A.* **101**:4758-4763.

Steketee, K., Ziel-van der Made, A.C., van der Korput, H.A., Houtsmuller A.B. and Trapman J. (2004) A bioinformatics-based functional analysis shows that the specifically androgen-regulated

gene SARG contains an active direct repeat androgen response element in the first intron. J. *Mol. Endocrinol.* **33**, 477-491

Tanner, T., Claessens, F., Haelens, A. (2004) The hinge region of the androgen receptor plays a role in proteasome-mediated transcriptional activation. *Ann. N.Y.Acad. Sci.* **1030**, 586-590

Verrijdt, G., Schoenmakers, E., Haelens, A., Peeters, B., Verhoeven, G., Rombauts, W. and Claessens F. (2000) Change of specificity mutations in androgen-selective enhancers: evidence for a role of differential DNA-binding by the androgen receptor. *J. Biol. Chem.* **275**: 12298-12305

Wang Q, Lu J, Yong EL (2001) Ligand- and coactivator-mediated transactivation function (AF2) of the androgen receptor ligand-binding domain is inhibited by the cognate hinge region. *J. Biol. Chem*.**276**:7493-7499.

Wong, H.Y., Burghoorn, J.A., van Leeuwen, M., De Ruiter P.E., Schippers, E., Blok, L.J., Dekker, H.L., De Jong, L., Trapman, J., Grootegoed, J.A., Brinkmann, A.O. (2004) *Biochem. J.* **383**, 267-76.

Zhou, Z.X., Kemppainen, J.A., Wilson, E.M. (1995) Identification of three proline-directed phosphorylation sites in the human androgen receptor. *Mol. Endocrinol.* **9**, 605-615.

Structural basis of androgen receptor binding to selective androgen response elements

Paul L. Shaffer*†, Arif Jivan*†, D. Eric Dollins*, Frank Claessens‡, and Daniel T. Gewirth*§

*Department of Biochemistry, Duke University Medical Center, Durham, NC 27710; and ‡Division of Biochemistry, Faculty of Medicine, Campus Gasthuisberg, University of Leuven, 3000 Leuven, Belgium

Communicated by Jane S. Richardson, Duke University Medical Center, Durham, NC, February 17, 2004 (received for review October 29, 2003)

Steroid receptors bind as dimers to a degenerate set of response elements containing inverted repeats of a hexameric half-site separated by 3 bp of spacer (IR3). Naturally occurring selective androgen response elements have recently been identified that resemble direct repeats of the hexameric half-site (ADR3). The 3D crystal structure of the androgen receptor (AR) DNA-binding domain bound to a selective ADR3 reveals an unexpected head-to-head arrangement of the two protomers rather than the expected head-to-tail arrangement seen in nuclear receptors bound to response elements of similar geometry. Compared with the glucocorticoid receptor, the DNA-binding domain dimer interface of the AR has additional interactions that stabilize the AR dimer and increase the affinity for nonconsensus response elements. This increased interfacial stability compared with the other steroid receptors may account for the selective binding of AR to ADR3 response elements.

The androgen receptor (AR) is a ligand-activated transcription factor that plays a central role in male sexual development and in the etiology of prostate cancer (1, 2). It is a member of the steroid and nuclear hormone receptor superfamily, which also includes receptors for glucocorticoids (GR), mineralocorticoids (MR), progesterone (PR), estrogens (ER), and vitamin D (VDR) (3). Members of this family contain conserved, discrete, DNA-binding domains (DBDs) and ligand-binding domains. The amino-terminal domain and the hinge region connecting the central DBD to the C-terminal ligand-binding domain diverge among family members.

The hormone receptor DBD consists of a highly conserved 66-residue core made up of two zinc-nucleated modules, shown schematically in Fig. 1*A* (4, 5). With VDR as the only reported exception (6), the isolated DBD and associated C-terminal extension are necessary and sufficient to generate the same pattern of DNA response element selectivity, partner selection, and dimerization as the full-length receptor from which it is derived (6–11).

Although ligand binding elicits distinct hormone-specific responses, all classical steroid receptors (AR, PR, MR, and GR) recognize identical DNA response elements, which consist of two hexameric half-sites (5′-AGAACA-3′) arranged as inverted repeats with 3 bp of separating DNA, producing the 2-fold IR3 sequence pattern (Fig. 1*B*) (12). A question that continues to engage the steroid receptor field is how these transcription factors achieve DNA target specificity despite this degeneracy. As seen in the structures of the GR and ER DBDs bound to IR3 elements (4, 13), the receptors bind as "head-to-head" homodimers whose symmetric displacement across the DNA pseudodyad reflects the underlying half-site arrangement. Differences in steroid metabolism, receptor expression, local chromatin structure, and the availability of cofactors all contribute to steroid-specific responses (14–17). However, recent work has now also identified selective androgen response elements (AREs). The AREs consist of two hexameric half-sites arranged as an androgen direct repeat separated by 3 bp of spacer (ADR3) (18–21), with the half-site repeating on the same strand (Fig. 1*B*). The expanded binding repertoire of AR, including both the common IR3 and specific ADR3 elements, breaks the degeneracy of the steroid response elements, allowing specific AR activation from certain response elements but disfavoring interaction with PR, MR, or GR. This finding could further account for steroid-specific actions *in vivo*.

The crystal structures of nuclear receptors bound to direct-repeat elements, including the VDR DBD bound to a similar DR3 element, reveal a "head-to-tail" protein dimer bound to the DNA (6, 22–24). For AR to bind to ADR3-type elements in a head-to-tail orientation, the DBD would require a second dimerization interface that is distinct from the canonical D box region used to dimerize on IR3 elements (25). To visualize this unusual homodimeric assembly, we have solved the crystal structure of an AR DBD homodimer bound to an ADR3 response element. The structure we report here reveals that the proteins do not adopt the expected head-to-tail orientation on the DNA, but, instead, they retain the symmetric mode of dimerization observed previously for the GR DBD bound to an IR3 DNA element. We describe the protein–protein and protein–DNA interactions that allow for this unexpected arrangement, and we propose that AR-specific dimerization contacts account for the AR specificity of ADR3 elements.

Materials and Methods

Protein and DNA Purification. The rat AR DBD (residues 533–637, C552A) was expressed in *Escherichia coli* BL21/DE3 cells as a GST fusion and purified with a glutathione-Sepharose column (Sigma). The GST was cleaved with thrombin at 4°C overnight. Further purification was performed with SP Sepharose FastFlow (pH 7.4) and Source 15S (pH 6.9) columns. Protein concentration and purity was determined by UV absorbance and SDS/PAGE.

Synthetic oligonucleotides (W. M. Keck Facility, Yale University) were detritylated and purified by reversed-phase HPLC (Rainin Dynamax-300). Concentrated, purified strands were annealed by heating to 95°C and slowly cooling to room temperature.

Crystallization and Data Collection. Samples for cocrystallization contained DNA and protein concentrations of 0.15 and 0.30 mM, respectively, in 5 mM Tris (pH 7.6)/150 mM LiCl/10 mM DTT. Crystals were grown by hanging drop vapor diffusion at 18°C with the addition of 2 μl of the complex to an equal volume of reservoir solution (50 mM Mes, pH 5.6/0–20 mM $MgCl_2$/0–2% polyethylene glycol 400). Diffraction quality crystals (0.15 × 0.15 × 0.4 mm) grew in 2–6 weeks.

Crystals were equilibrated into reservoir solution supplemented with 35% glycerol before being flash-cooled in liquid

Abbreviations: AR, androgen receptor; GR, glucocorticoid receptor; MR, mineralocorticoid receptor; PR, progesterone receptor; ER, estrogen receptor; VDR, vitamin D receptor; DBD, DNA-binding domain; ARE, androgen response element.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1R4I).

†P.L.S. and A.J. contributed equally to this work.

§To whom correspondence should be addressed. E-mail: gewirth@duke.edu.

www.pnas.org/cgi/doi/10.1073/pnas.0401123101

nitrogen. Diffraction data were collected at −180°C on beamline 22ID at the Advanced Photon Source with a CCD detector (Marresearch, Norderstedt, Germany). Data were indexed and reduced by using HKL2000 (26).

Structure Determination and Refinement. Four zinc sites were found by using SOLVE (27) and data from the peak anomalous wavelength. Experimental phases were generated with these sites; and, in the anomalous difference Fourier maps, the four zinc sites had peaks of >30 σ, whereas the next highest peak was 3 σ, indicating one AR dimer was in the asymmetric unit. Only one of the two possible enantiomeric space group choices yielded zinc sites that corresponded to possible AR dimers. Visual inspection of the zinc sites revealed that the proteins were arranged in a palindromic orientation. This finding led to construction of a molecular replacement model by using the ER DBD-IR3 structure (13) (PDB ID code 1HCQ). Because of its higher sequence homology to AR, the ER DBD was replaced with the core GR DBD (4) (PDB ID code 1GLU) by using least-squares fitting. A molecular replacement solution was obtained by using MOLREP (28).

Multiwavelength anomalous dispersion phases were calculated by using the remote and peak wavelength data to 3.4 Å and also used in refinement, which was done in CNS (29) by using the maximum likelihood Hendrickson–Lattman target. Model building was done by using O (30). Even at 3.1 Å, the number of unique reflections used was eight times the number of modeled atoms because of the very large (>80%) solvent content of the crystal, allowing for restrained individual B factor refinement in later rounds. Visualization of hydrogen bonds, van der Waals interactions, and clashes was aided by use of all atom contacts in KING and PROBE (31). Graphics used RIBBONS (32) and PYMOL (DeLano Scientific, San Carlos, CA).

Results

Crystallization and Structure Solution. Initial crystals of AR DBD–ADR3 complexes grew as thin needles from complexes containing AR DBD (residues 533–619) and diffracted to 4 Å with synchrotron radiation. These crystals were resistant to dissolution, suggesting crosslinking within the lattice. The AR DBD contains a nonconserved cysteine at position 552[11] (common receptor DBD numbering is given in brackets), which was predicted to be solvent-exposed based on modeling from the GR DBD structure. When Cys-552[11] in the AR DBD was changed to alanine, complexes containing this mutant yielded bar-shaped crystals that were isomorphous with the initial crystal form. These crystals were used to determine the structure of the AR DBD–DNA complex (PDB ID code 1R4I).

The structure of AR DBD(533–637)Cys552Ala in complex with ADR3 DNA (Fig. 1) was determined at 3.1 Å by a combined MAD and molecular replacement approach with diffraction data collected at the zinc anomalous edge. The arrangement of the proteins on the ADR3 DNA was determined from zinc anomalous data that revealed the location of the four zinc atoms in the complex. Data collection and refinement statistics are presented in Table 1, and representative electron density maps are shown Fig. 7, which is published as supporting information on the PNAS web site.

Anomalous difference Fourier maps confirmed that the asymmetric unit consists of just one AR DBD homodimer–DNA complex, yielding a Matthews number of 6.9 and a solvent content of 82%. The main crystal-packing interactions are made by the junction near protomer A, which contains neither a pseudocontinuous DNA interaction nor a biologically plausible alternative protein dimer interface. The downstream AR DBD (protomer B) makes only two crystal contacts by residues Phe-589[48] and Arg-590[49] and, except for the interaction with protomer A and the DNA, it is otherwise completely exposed to the large solvent channels (Fig. 2).

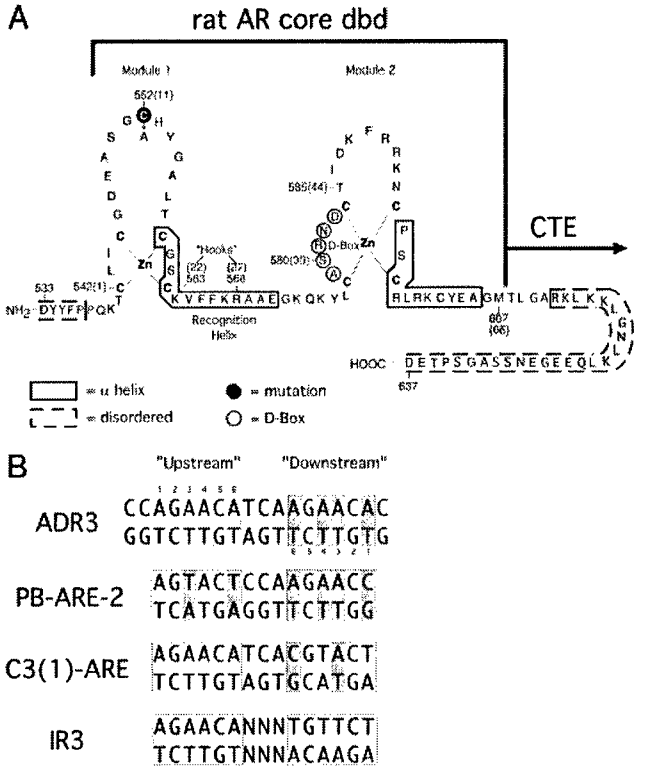


Fig. 1. Protein and DNA constructs. (*A*) The rat AR DBD. Sequence numbers in parentheses refer to the common receptor DBD-numbering scheme. Residues in dashed boxes are disordered in both protomers of the homodimeric complex. (*B*) The DNA used in cocrystallization, labeled ADR3, two naturally occurring AR response elements, PB-ARE-2 and C3 (1)-ARE, and a canonical IR3 steroid response element. Differences from the IR3 sequence are shaded gray.

Examination of the crystal-packing interactions can explain the refractory effect of C552[11] on crystallization. Residue

Table 1. Summary of data collection and refinement

Diffraction data		
Space group,*† Å	$P3_221$ 137.89, 85.71	
Data set	Native/remote	Zn peak
Wavelength, Å	1.0000	1.2831
Resolution, Å	50–3.1	50–3.4
Last shell, Å	3.21–3.1	3.52–3.4
Unique reflections	17,313	25,060
Completeness, % (last shell)	99.7 (99.1)	99.5 (98.9)
Average I/σ_I (last shell)	20.4 (2.5)	19.6 (2.0)
R_{merge}, % (last shell)	9.6 (62)	7.8 (58)
FOM (after DM)‡	0.41 (0.96)	
Crystallographic refinement		
Resolution range, Å	50–3.1	
Reflections ($F > 2\sigma_F$)	14,839 (12,418)	
Atoms	1,813	
rms bond lengths, Å	0.0076	
rms bond angles, °	1.29	
R value ($F > 2\sigma_F$)†	24.6 (22.7)	
R_{free} ($F > 2\sigma_F$)	26.4 (24.9)	

*$R_{merge} = \Sigma_{hkl}\Sigma_i|I_{i(hkl)} - \langle I_{(hkl)}\rangle|/\Sigma_{hkl}\Sigma_i I_{(hkl)}$.
†$R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$. 5% of the reflections were used for R_{free}.
‡Figure of merit = $\langle|\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)|\rangle$, where α is the phase and $P(\alpha)$ is the phase-probability distribution.

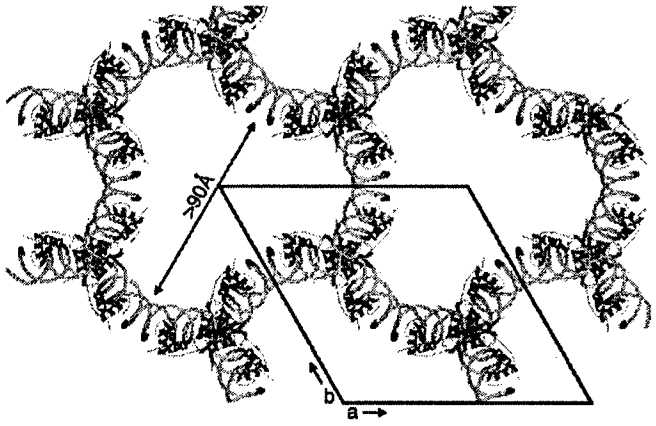


Fig. 2. Crystal packing of the AR DBD–ADR3 complex. Red and blue ribbons are the upstream and downstream subunits, respectively, with the DNA backbone shown in gold. The view is parallel to the *c* axis of the crystal, and the unit cell is shown.

552[11] from protomer A is in position to crosslink with Cys-578[37] of protomer A in the adjacent symmetry-related complex. Cys-578[37] coordinates a zinc atom in the first Zn module. Formation of a C552[11]-C578[37] disulfide link is likely to disrupt the native AR DBD conformation and adversely affect crystal order.

The AR DBDs Are Arranged as an Inverted Repeat on a Direct-Repeat DNA Target. In all the dimeric hormone receptor DBD–DNA complexes determined to date, the two DBDs adopt the same relative orientation as that of the underlying DNA target. Surprisingly, however, in the structure of AR DBD bound to ADR3 DNA, the two AR DBD protomers are not arranged as a head-to-tail dimer, as would be expected of receptors bound to a direct-repeat DNA element. Instead, the proteins form a symmetric, head-to-head dimer that is nearly identical with the dimer seen in the ER DBD–DNA and GR DBD–DNA structures (rms deviation for α-carbons of 1.09 and 0.89 Å, respectively) (4, 13). This finding was confirmed unambiguously by inspection of the positions of the four zinc sites determined from anomalous difference maps calculated from single wavelength anomalous dispersion phases (Fig. 3). The arrangement of the AR dimer is unlikely to be an artifact of crystal packing, because there are only two small crystal contacts between the downstream DBD (protomer B) and the neighboring molecules in the crystal lattice (Fig. 2).

The AR DBD Homodimer Interface. The subunit interface of the AR DBD homodimer is symmetric and closely resembles that seen in the GR DBD–DNA complex (4). As in the GR DBD– and ER DBD–DNA complexes, the majority of the cross-subunit contacts are made in the D box region of the second zinc module. In the GR homodimer, the subunit interface is stabilized both by a network of hydrogen bonds between D box residues and by an extensive complementary surface. As seen in Fig. 4*B*, however, the GR interface contains a void formed where the Gly-478[39] from the opposing subunits face each other. This "glycine hole" is also a feature of the MR and PR. In the AR DBD, however, glycine is replaced by Ser-580[39]. This serine packs into the glycine hole of the dimer interface, filling the void and making van der Waals contact with its counterpart in the other subunit. In addition, the arrangement of the two serines is optimal for the formation of a hydrogen bond across the molecular pseudodyad. The substitution of serine for glycine in the AR D box is likely

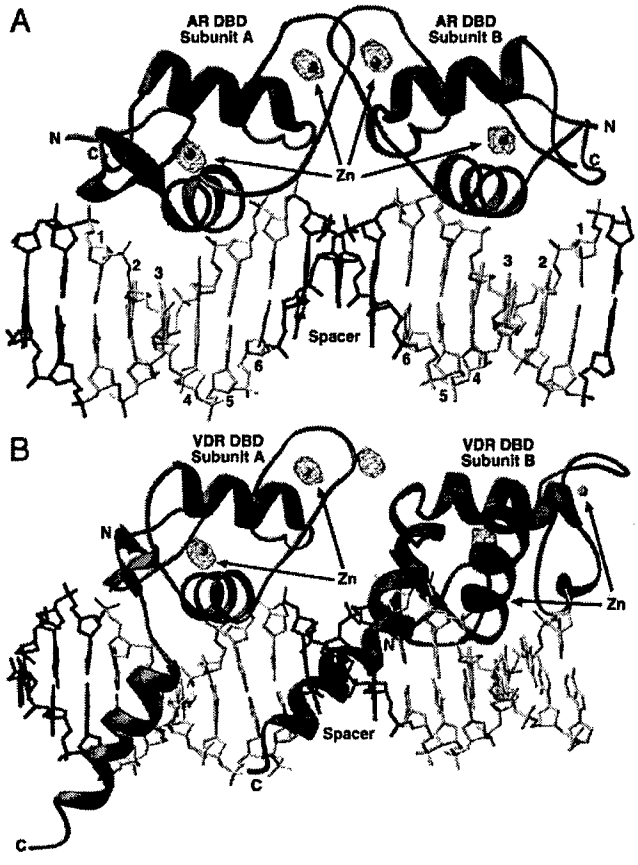


Fig. 3. Overall architecture of the AR DBD–ADR3 and VDR DBD–DR3 complexes. (*A*) The AR DBD–ADR3 complex. The two protomers are in red and blue, the hexameric half-site DNA is gold, and the spacer and flanking base pairs are black. In brown is a 20-σ contour of the experimental anomalous Fourier difference map. (*B*) The VDR DBD–DR3 complex. VDR DBD protomer A is shown in the same orientation as the AR DBD subunit A in *A*. The zincs of subunit B fail to occupy the peaks in the anomalous difference Fourier map in this dimeric arrangement, indicating the AR DBD does not form a head-to-tail dimer.

to increase the relative strength of the dimer interface of the AR DBD.

The AR DBD also makes an additional pair of symmetrical contacts between Thr-585[44] and the carbonyl oxygen of Ala-579[38] in the opposing protomer. In the GR DBD the residue at this position is an isoleucine, and replacement with a threonine as seen in the AR is likely to increase the stability of the dimer because of the enthalpic contribution of the additional two hydrogen bonds. In addition, the change from Ile in GR to Thr in AR removes a nonpolar residue from the solvent-exposed surface of the DBD, thus entropically stabilizing the AR as well.

The AR DBD (P.L.S. and D.T.G., unpublished work) and GR DBD (33) are monomers in solution. Because cooperative dimerization greatly increases the affinity of receptors for their bipartite response elements, these two changes should also increase the relative affinity of the AR for a given response element compared with GR. In support of this hypothesis, GR DBD mutants containing a serine in place of Gly-478[39] in the D box or a threonine in place of GR Ile-483[44] show increased affinity for both palindromic and direct-repeat response elements compared with wild type (34), confirming the importance of these interactions for dimer stability.

Protein–DNA Interactions. The DNA used for cocrystallization has a DR3 arrangement of hexameric half-sites, with the sense strand

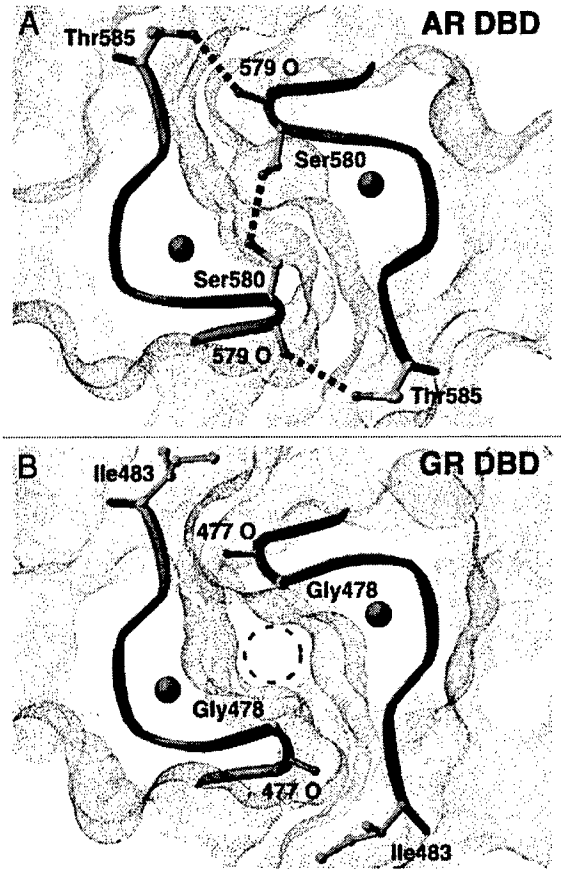


Fig. 4. (*A*) The AR DBD dimer interface. The molecular surfaces of the AR subunits are shown in red and blue. Dashed black lines are hydrogen bonds. (*B*) A similar view of the GR DBD dimer interface. The "glycine hole" is noted by the dashed circle.

sequence 5′-CC **AGAACA** TCA **AGAACA** G-3′. However, the AR proteins were observed to bind in a symmetric, head-to-head arrangement, as was seen with steroid receptors bound to an IR3 response element (symmetrized consensus sequence of 5′-**AGAACA** NNN **TGTTCT**-3′). One half-site, bound by protomer A and shown here as upstream, is common to both DR3 and IR3 elements and is a high-affinity, consensus-binding site for steroid DBDs. Protomer B, on the other hand, binds to the downstream half-site that contains the consensus IR3-type bases at only the second and fifth positions. Experimentally phased electron density maps were used to identify the length of the asymmetric flanking sequences and unambiguously assign the orientation of the DNA. Within the limitations imposed by the diffraction resolution, the DNA does not exhibit significant deviations from B form.

Backbone DNA contacts are similar for both AR protomers (Fig. 5) and show the pattern seen previously in structures of steroid receptor–DNA complexes (4, 35). The base-specific contacts between the AR DBD and the consensus half-site are also nearly identical with those of the GR DBD to its cognate half-site and are shown in Fig. 5*A*. In addition to these previously described interactions, we also note that the aliphatic portion of the Arg-568[27] side chain makes additional van der Waals contacts with Val-564[23] and the C5 methyl group of the thymine at the sixth position of the consensus half-site. Thymine is the only base that can form the second half of this van der Waals "sandwich," and this specific contact likely explains why an A:T base pair is commonly observed at the sixth position of AR-specific half-sites (Fig. 6). Because the interaction between the conserved arginine and thymine is also present in consensus half-sites in the GR, ER, 9-*cis*-retinoic acid receptor, and other steroid and nuclear hormone receptor DBD structures, this can explain the preference for the A:T base pair at the sixth position in these protein–DNA complexes as well.

The nonconsensus half-site interaction seen in the AR DBD–ADR3 structure contains the top strand sequence 5′-AGAACA-3′, with the two bases that match the consensus for a downstream IR3 half-site underlined. These two bases lie at the correct IR3 positions because they are symmetric within the hexameric half-site. This serendipitous match to the consensus IR3 half-site allows Lys-563[22] and Arg-568[27] of protomer B to recapitulate the hydrogen bonds to the GC base pairs at positions 2 and 5 of the hexameric half-site, as seen in the upstream element. These two "hooks" are common elements that position the recognition helix within the major groove of the hexameric half-site (36).

In the cognate AR DBD half-complex, the side chain of Val-564[23] makes van der Waals contact with the 5-methyl group of the T4 of the antisense strand. This interaction between the two nonpolar substituents is the discriminating feature of specific steroid receptor–DNA interfaces, and the resulting dehydration of the protein–DNA interface contributes entropic stabilization to the binding (35, 37). In the nonconsensus AR half-complex, A replaces the T at position 4 of the sense strand, resulting in the loss of the Val-564[23]-T4 contact. Although this replacement reduces the number of specific, stabilizing, interactions with the DNA half-site, the substitution of an A base for the consensus T does not cause a steric clash that might disfavor binding to this element. As befits the reduced complementarity between the AR DBD and the nonconsensus half-site, the cognate half-complex buries slightly more surface area from solvent (1,230 Å^2) than the noncognate one (960 Å^2).

AR Mutations. Mutations in the AR DBD associated with partial or complete androgen insensitivity (see ww2.mcgill.ca/androgendb) can be understood mechanistically in light of the structure determined here. Many of these were correctly analyzed earlier based on the structure of the GR DBD (38). More recently, within the D box, Ala579Thr (39–41) and Ser580Thr (42) mutations have been reported to lead to loss of AR dimerization. Modeling the Ser580Thr mutation on the AR DBD dimer leads to bad steric clashes in any possible Thr conformation, forcing backbone shifts that presumably disfavor dimerization. Modeling of the Ala579Thr substitution is more problematic, because the Thr side chains can each be accommodated with modest steric overlaps of 0.3–0.4 Å. However, that may be enough to force structural changes in the interface, and the imprecision of low resolution may underestimate the problem. The Ala579Thr mutation can be relieved by a compensatory change in Thr-585 to Ala (43), close to residue 579 across the dimer interface. This further change may relieve strains in the dimer interface or in the Zn ligand geometry caused by the Ala579Thr mutation.

Discussion

We have determined the structure of the AR DBD bound to an idealized steroid DR3 response element. Based on studies of the VDR DBD (6), which also binds to a DR3-type response element, we expected the tandem arrangement of half-sites to direct head-to-tail binding of the AR DBD to the DNA. Surprisingly, however, the AR DBDs bind to the direct-repeat response element as head-to-head symmetrical dimers. This mismatch between receptor dimer- and response element-arrangement results in one AR DBD bound to a high-affinity cognate half-site, and the partner DBD bound to a lower-affinity half-site. This finding indicates that the energetic penalty in-

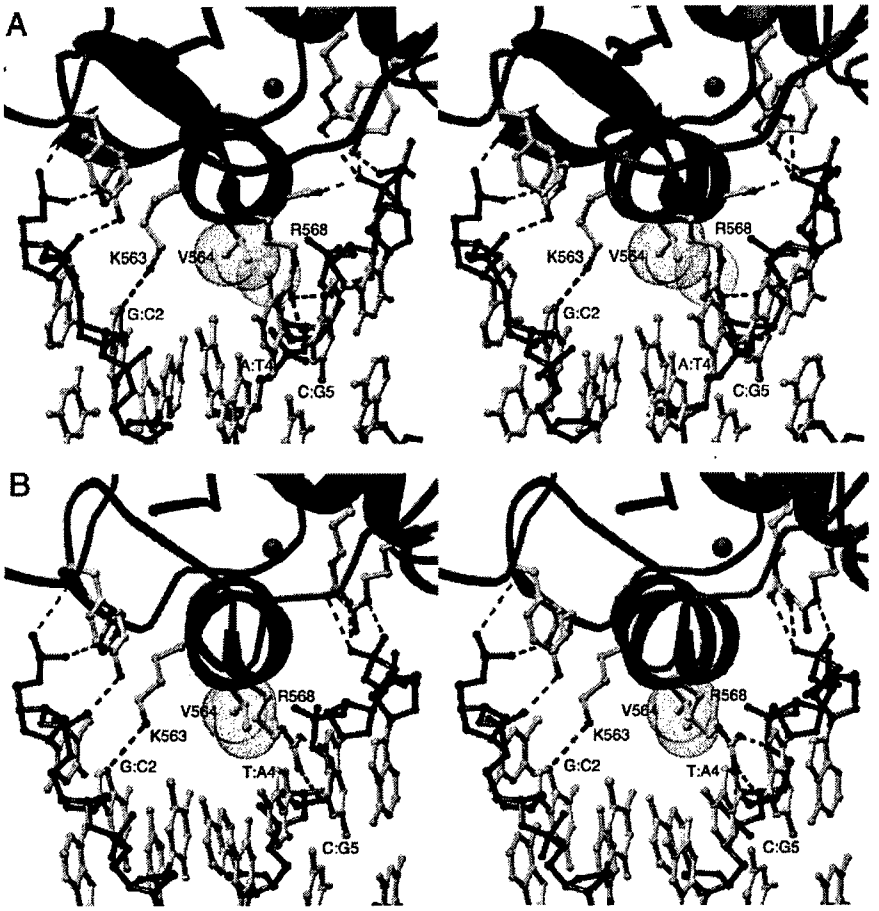


Fig. 5. Stereoview of the AR DBD–DNA interfaces. (*A*) The upstream, cognate, protein–DNA interface. (*B*) The downstream, noncognate interface. The protein is shown in the same orientation as in *A*.

curred by binding to a less favored half-site sequence is more than offset by maintaining the preferred IR3-type dimer interface. This finding is analogous to an earlier observation that the GR DBD maintains the IR3 dimer interface and spacing even when challenged with an IR4 response element (4).

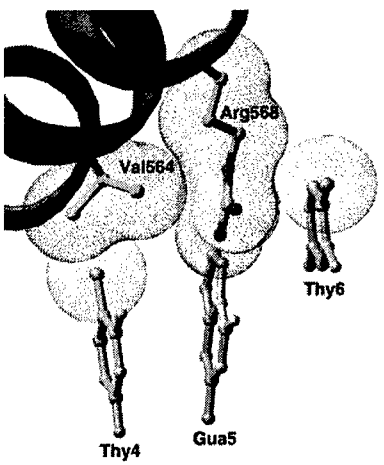


Fig. 6. The arginine "sandwich." Val-564 and Arg-568 of the AR DBD subunit along with bases T4, G5, and T6 of the antisense strand of the upstream, cognate half-site are shown. The C5 methyl group of T6 forms van der Waals interactions with one face of Arg-568, whereas the other side packs against Val-564.

Both the AR and the GR exhibit similar interactions with steroid response elements, yet the AR exhibits consistently stronger binding to direct repeat-type response elements than does the GR. Some of this difference in affinity may be attributable to differences in the C-terminal extension of each DBD, although in both GR and AR these regions were disordered in the crystal structure and may contribute only general electrostatic interactions without affecting selectivity or discrimination. Within the core of the DBD, however, the protein–DNA interactions are nearly identical for both receptor DBDs, and much of the difference in response element affinity is therefore likely to reside in the ability of each receptor to cooperatively form head-to-head dimers on bipartite response elements where the interaction with one or both hexameric half-sites is nonoptimal.

The second zinc module has been shown to be necessary for AR to bind cooperatively to ADR3s (44). The steroid receptor DBD dimerization interface is contained within this module, and between AR and GR it differs at just four positions. The increased AR dimer affinity can be explained by two of these four substitutions, one in the D box, and the other two residues beyond. In the D box, AR is the only steroid receptor that has a Ser residue at the second position, Ser-580[39], and this serine packs into the core of the dimer interface, making both van der Waals interactions and a cross-subunit hydrogen bond. All other steroid receptors have a Gly at this position, which lacks this additional hydrogen bond and leaves a void in the interface. Two residues beyond the D box, an Ile-to-Thr substitution in AR allows both a favorable cross-subunit side chain-to-backbone

hydrogen bond and removes the nonpolar Ile side chain from exposure to solvent. Together these two substitutions appear to account for the stronger AR dimer interface. These substitutions in turn allow the receptor to bind to a more diverse set of response elements with higher affinity and cooperativity than the GR.

Biochemical evidence for the increased cooperativity of the AR DBD dimer correlates with these structural observations. All the steroid receptors (MR, PR, GR, and AR) show a 5- to 10-fold lower affinity for the naturally occurring PB-ARE-2 DR3-type element than the C3 (1) IR3-type element (34). However, the AR DBD binds 3- to 10-fold better to both elements relative to the other steroid receptors. Thus, the binding constant for AR on an apparent DR3 target (23 ± 5 nM) is the same as that of the other receptors for the more optimal IR3 element (the average of the other three is 23 ± 9 nM) (44). Because the concentration of individual steroid receptors in the cell is approximately nanomolar, differences in binding constants of this order are likely to be significant. AR substitutions in the GR dimerization interface, including Gly483Ser and Ile483Thr, show higher affinity binding to both DR3 and IR3 response elements (34), thus mimicking the behavior of the AR. Together with the structural data, these observations suggest a model where, because of the increased strength of the AR dimer interface, AR-selective gene activation arises from the ability of the AR to bind to IR3 response elements that have a greater deviation from the consensus half-site sequence. The reverse cross-activation of GR-responsive genes by the AR would likely be disfavored by the highly tissue-specific expression pattern of the AR compared with the GR.

The structure of the AR DBD bound as an inverted repeat to a direct-repeat response element highlights the fact that DNA target recognition by hormone receptors is strongly governed by the dimerization behavior of the two interacting protomers, even at the cost of losing specific interactions with the target DNA. With the exception of the Ecdysone receptor, which binds to IR1 rather than IR3 targets consisting of AGGTCA rather than AGAACA half-sites (45), no physiologically relevant dimerization interface within the classical steroid receptor DBDs, other than the primary one, has been observed to date in structural studies. Moreover, attempts to capture such potential alternative interfaces, as described in this report, and previously for GR (4), have been unfruitful. This in turn implies that selective hormone response elements that appear to have alternative arrangements of their hexameric half-sites, such as the pemARE with a proposed 5-bp spacer between half-sites (46), may instead simply be further examples of the ability of these receptors to exploit the strength of their DBD dimerization interfaces to accommodate suboptimal protein–half-site interactions. This ability is likely to be not only a mechanism of response element discrimination, but also an effective way of modulating transcription from different hormone-responsive genes.

We thank Nikki Fetter and Jenna Vanliere for help with crystallization and Karen Soldano for preparative assistance. This work was supported by U.S. Army Prostate Cancer Research Program grants (to D.T.G. and F.C.).

1. McPhaul, M. J. (1999) *J. Steroid Biochem. Mol. Biol.* **69,** 315–322.
2. Gottlieb, B., Lehvaslaiho, H., Beitel, L. K., Lumbroso, R., Pinsky, L. & Trifiro, M. (1998) *Nucleic Acids Res.* **26,** 234–238.
3. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., *et al.* (1995) *Cell* **83,** 835–839.
4. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R. & Sigler, P. B. (1991) *Nature* **352,** 497–505.
5. Khorasanizadeh, S. & Rastinejad, F. (2001) *Trends Biochem. Sci.* **26,** 384–390.
6. Shaffer, P. L. & Gewirth, D. T. (2002) *EMBO J.* **21,** 2242–2252.
7. Mader, S., Chen, J. Y., Chen, Z., White, J., Chambon, P. & Gronemeyer, H. (1993) *EMBO J.* **12,** 5029–5041.
8. Perlmann, T., Rangarajan, P. N., Umesono, K. & Evans, R. M. (1993) *Genes Dev.* **7,** 1411–1422.
9. Towers, T. L., Luisi, B. F., Asianov, A. & Freedman, L. P. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6310–6314.
10. Zechel, C., Shen, X. Q., Chambon, P. & Gronemeyer, H. (1994) *EMBO J.* **13,** 1414–1424.
11. Zechel, C., Shen, X. Q., Chen, J. Y., Chen, Z. P., Chambon, P. & Gronemeyer, H. (1994) *EMBO J.* **13,** 1425–1433.
12. Beato, M., Herrlich, P. & Schutz, G. (1995) *Cell* **83,** 851–857.
13. Schwabe, J. W., Chapman, L., Finch, J. T. & Rhodes, D. (1993) *Cell* **75,** 567–578.
14. Strahle, U., Boshart, M., Klock, G., Stewart, F. & Schutz, G. (1989) *Nature* **339,** 629–632.
15. Funder, J. W. (1993) *Science* **259,** 1132–1133.
16. Muller, J. M., Isele, U., Metzger, E., Rempel, A., Moser, M., Pscherer, A., Breyer, T., Holubarsch, C., Buettner, R. & Schule, R. (2000) *EMBO J.* **19,** 359–369.
17. List, H. J., Lozano, C., Lu, J., Danielsen, M., Wellstein, A. & Riegel, A. T. (1999) *Exp. Cell Res.* **250,** 414–422.
18. Claessens, F., Alen, P., Devos, A., Peeters, B., Verhoeven, G. & Rombauts, W. (1996) *J. Biol. Chem.* **271,** 19013–19016.
19. Rennie, P. S., Bruchovsky, N., Leco, K. J., Sheppard, P. C., McQueen, S. A., Cheng, H., Snoek, R., Hamel, A., Bock, M. E. & MacDonald, B. S. (1993) *Mol. Endocrinol.* **7,** 23–36.
20. Verrijdt, G., Schoenmakers, E., Alen, P., Haelens, A., Peeters, B., Rombauts, W. & Claessens, F. (1999) *Mol. Endocrinol.* **13,** 1558–1570.
21. Verrijdt, G., Schoenmakers, E., Haelens, A., Peeters, B., Verhoeven, G., Rombauts, W. & Claessens, F. (2000) *J. Biol. Chem.* **275,** 12298–12305.
22. Rastinejad, F., Perlmann, T., Evans, R. M. & Sigler, P. B. (1995) *Nature* **375,** 203–211.
23. Rastinejad, F., Wagner, T., Zhao, Q. & Khorasanizadeh, S. (2000) *EMBO J.* **19,** 1045–1054.
24. Zhao, Q., Chasse, S. A., Devarakonda, S., Sierk, M. L., Ahvazi, B. & Rastinejad, F. (2000) *J. Mol. Biol.* **296,** 509–520.
25. Verrijdt, G., Haelens, A. & Claessens, F. (2003) *Mol. Genet. Metab.* **78,** 175–185.
26. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
27. Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55,** 849–861.
28. Vagin, A. A. & Isupov, M. N. (2001) *Acta Crystallogr. D.* **57,** 1451–1456.
29. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905–921.
30. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110–119.
31. Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S. & Richardson, D. C. (1999) *J. Mol. Biol.* **285,** 1711–1733.
32. Carson, M. (1991) *J. Appl. Crystallogr.* **24,** 958–961.
33. Freedman, L. P., Yamamoto, K. R., Luisi, B. F. & Sigler, P. B. (1988) *Cell* **54,** 444 (lett.).
34. Schoenmakers, E., Verrijdt, G., Peeters, B., Verhoeven, G., Rombauts, W. & Claessens, F. (2000) *J. Biol. Chem.* **275,** 12290–12297.
35. Gewirth, D. T. & Sigler, P. B. (1995) *Nat. Struct. Biol.* **2,** 386–394.
36. Nelson, C. C., Hendy, S. C., Shukin, R. J., Cheng, H., Bruchovsky, N., Koop, B. F. & Rennie, P. S. (1999) *Mol. Endocrinol.* **13,** 2090–2107.
37. Lundback, T., Cairns, C., Gustafsson, J. A., Carlstedt-Duke, J. & Hard, T. (1993) *Biochemistry* **32,** 5074–5082.
38. Luisi, B. F., Schwabe, J. W. & Freedman, L. P. (1994) in *Vitamins and Hormones*, ed. Litwack, G. (Academic, New York), Vol. 49, pp. 1–47.
39. Gast, A., Neuschmid-Kaspar, F., Klocker, H. & Cato, A. C. (1995) *Mol. Cell Endocrinol.* **111,** 93–98.
40. Holterhus, P. M., Wiebel, J., Sinnecker, G. H., Bruggenwirth, H. T., Sippell, W. G., Brinkmann, A. O., Kruse, K. & Hiort, O. (1999) *Pediatr. Res.* **46,** 684–690.
41. Lundberg Giwercman, Y., Nikoshkov, A., Lindsten, K., Bystrom, B., Pousette, A., Knudtzon, J., Alm, J. & Wedell, A. (2000) *Horm. Res.* **53,** 83–88.
42. Nordenskjold, A., Friedman, E., Tapper-Persson, M., Soderhall, C., Leviav, A., Svensson, J. & Anvret, M. (1999) *Urol. Res.* **27,** 49–55.
43. Kaspar, F., Klocker, H., Denninger, A. & Cato, A. C. (1993) *Mol. Cell. Biol.* **13,** 7850–7858.
44. Schoenmakers, E., Alen, P., Verrijdt, G., Peeters, B., Verhoeven, G., Rombauts, W. & Claessens, F. (1999) *Biochem. J.* **341,** 515–521.
45. Devarakonda, S., Harp, J. M., Kim, Y., Ozyhar, A. & Rastinejad, F. (2003) *EMBO J.* **22,** 5827–5840.
46. Geserick, C., Meyer, H. A., Barbulescu, K. & Haendler, B. (2003) *Mol. Endocrinol.* **17,** 1738–1750.